UNIVERSITY OF EDINBURGH

'AGE AND COLOUR VISION'

by

R. LAKOWSKI, M.A. Dip Ed.

Thesis submitted for the degree of Doctor of Philosophy in the faculty of Social Sciences.

October, 1964.

AGE AND COLOUR VISION

Table of Contents

Acknowledgements

The writer gratefully acknowledges the helpful advice and encouragement given to him by Dr. Collins during the preparation of this thesis, and thanks Prof. Drever for his advice and help in securing the necessary facilities and equipment for some of these experiments.

He further acknowledges the skilled assistance of Mr. J. Rae, the laboratory's artist, in the preparation of the numerous diagrams. His thanks are extended to the other members of the visual laboratory, especially to Mr. P.K. Kinnear, B.Sc., and Mr. J. Gill, B.Sc., for comments and help with some of the calculations for the statistical and colorimetric analysis. He wishes to thank his wife for reading and correcting the various drafts of the thesis, and would also like to acknowledge the excellent work done by Mrs. E. Smith in typing the final manuscript, in the very short time available.

Lastly, he would like to acknowledge the help of Professor Pickford of Glasgow University for first arousing his interest in the subject matter of this thesis, and for making the anomaloscope available for this study.

1. INTRODUCTION TO THESIS

Testing for this study of age and colour vision was begun in 1955 and nearly completed by 1958. It was in that year that application for admission as a Ph. D. Student of Edinburgh University was made. Although the basic facts of this study collected from 1955-58 are embodied in this thesis, a great deal of material from further experimenting, analysing and testing is also included.

The main subjects of the thesis are :-

(A) A historical survey of previous studies on the problem of age and colour vision. This will deal with researches up till 1957 and will include an evaluation of the discrepancies between the various studies.

(B) A quantitative analysis of results obtained from the study for this thesis. This will include a discussion of performances on the various pseudo-isochromatic tests, and on the Pickford anomoloscope. In addition, a colorimetric and photometric study will be made of the tests used in the researches on age and colour vision. This will lead to an analysis of what is actually tested by the colour vision tests used by various research workers.

(C) Conclusions reached from this study and a comparison of the results of this work with the results of other researchers, both prior to 1957 and since then.

Much of the material presented here is completely new, and has not been published. Part of it, however, has appeared in print under the following titles, with the writer as sole author. In these papers (listed below) the subjec t matter presented is relevant to the main body of the thesis :-

1958 - 'Age and Colour Vision. The Advancement of Science '59' p. 231-237.

1959 - 'Experiments in Colour Vision, Especially with Reference to Ageing' Bull, Br. Psych. Soc. p. 38-39.

1961 - 'Is the deterioration of colour vision with age due to lens or retinal changes ? ' Tagungs Bericht International Farbtagung Dusseldorf 1961 - Teil 2 Die Farbe II Nr. 1/6 p. 69-89, 1962.

1962 - 'Testing of Colour Vision in Prospective Printers' Apprentices and the problems this Presents in Selection' - 6 emes Journes International de la Couleur, Evian. In print Br. J. of Physiological Optics and 'Couleurs' Revue Officielle du Centre D'Information de la Couleur, France.

In the historical part of this thesis, only studies which ended around 1957-58 are described. Any subsequent research is discussed in the section dealing with the main conclusions of this study, as evidence for or against any conclusions reached. The year 1958 was chosen as it marks the end of that period in which the problem of whether colour vision really does deteriorate with age, and at what point, were still open to question.

It should be noted here that in order to remain objective and unbiased, the writer did not acquaint himself with the research of others until after his own data had been collected.

2. HISTORICAL SURVEY

2 : 1 SURVEY OF PRE 1939 RESEARCH

Prior to 1940, age differences in colour discrimination were almost exclusively studied by those concerned with infant and child development and by colour theorists interested in ontogenetic explanations of colour discrimination.

Two types of experiment were attempted :-

1. experiments determining the age at which the ability to discriminate colour first appears, and
2. experiments which were aimed at finding out how colour discrimination developed.

Most of the experiments were of the first type. Psychologists such as McDougall (1906); Myers (1907); Valentine (1914); Pratt (1930) were all involved in colour experiments. The general conclusion reached by these authors was that colour discrimination occurs as early as the first month of life.

Sometimes, particularly in the earlier studies it is not clear whether care was taken to equate the brightnesses of the colours used with the result that the "colour" differences found might only be due to differences in brightness. However, careful experiments by Staples (1931-2) and Chase (1937) tend to nullify this criticism.

The growth of colour discrimination at a later stage of development was studied less often partly due to a lack of interest by these research workers and partly due to the difficulty of communicating in language, what is perceived. Binet (1890); Garbini (1894); and Monroe (1907) compared the relative abilities of pre-school children to match colours with their ability to name the same colour.

The results of Cook (1931) are typical of the uniform trend of this evidence. She found that, by the age of 2 years, children were able to match primary colours with an accuracy of 45% and to name them with an accuracy of 25%. By the age of 6 years, the two abilities have increased respectively to 97% and 62%. In several of the studies dealing with colour naming only, it was found that the red and blue were always named correctly, yellow and green less easily, and orange and violet were not completely known even at 14. Winch (1910) reported that various first colour words of children were used predominantly to distinguish chromatic from achromatic colours. As early as 1890, Wolfe found that black, white and red were the first colours to be named correctly by children and that they named green correctly more often than yellow. Luckiesh (1926) found that even the colour naming of adults is by no means constant! Hazlitt (1930) suggested that reports of poor colour discrimination up to the age of 4 may be due to a lack of interest rather than a lack of sensitivity and also to a failure to comprehend the task of sorting colours. Phillips (1938) said that intelligence was an important factor in colour discrimination.

The problem of age and colour vision began with Henry Clay Smith's work in the late thirties. (Though, Pierce (1934) reported a correlation of 0.51 between age and performance on a colour discrimination test, his results were based on data from only 40 female colour workers, and the age range was not specified.)

However, Joseph Tiffin and Hedwig Kuhn must share with Smith the honour of being pioneers of such studies, as their work, although it was started later, actually appeared in print one year earlier than Smith's, that is 1942.

Eight other studies relevant to our subject matter were carried out during the years between 1950-57. The exception was that of L. Boice whose work appeared in 1948.

These eight were :-

1) A. Chapanis' work published in 1950, on the relationship between age and visual acuity and colour vision. He used a battery of pseudo-isochromatic plates.
2) R. Brown's work at Glasgow, also in 1950 using the Ishihara plates and the Pickford Anomaloscope, on the colour vision of school children.
3) Robert Kleemeier in Florida who studied the relationship between the Ortho-Rater test of Acuity and Colour Vision in older people (1953).
4) F. L. Warburton's paper on the variations of colour vision in relation to practical colour matching, read in 1950 to the Colour Group. This appeared in print in 1953.
5) At Parma, Boles-Carenini, using the Nagel Anomaloscope, studied the problem of 'The chromatic sense in relation to age', in 1954.

6) Karla Janouskova's work in Prague on 'Age and Colour Vision' using the American Optical plates, (1950-55).

7) Louise Quellette's M. A. thesis at Fordham University on age differences in colour discrimination using the Color Aptitude Test, (1955).

8) Jeanne Gilbert's work on age changes in colour matching, again using the Color Aptitude Test, 1957.

Most of this historical survey will be confined to describing, occasionally in detail, the particulars of the above researches such as experimental design, the age and number of subjects, and the conclusions reached.

The chronological outline will be followed in order to show more fully the state of thought that was prevalent at that time. Originally studies were undertaken with the sole object of finding whether variations due to age existed as, for example, in the works of Smith and Tiffin and Kuhn, but succeeding studies except Gilbert's were carried out with the purpose either of confirming or discrediting the main hypothesis of Tiffin and Kuhn. This produced a confused picture of the effect of age on colour vision and in order to clarify the situation, a more detailed account of the works of Smith and Tiffin, Kuhn is necessary.

2 : 2 PIONEER STUDIES

(a) The Work of Henry Clay Smith

The first systematic study of the problem of progressive changes in colour discrimination from adolescence to senescence was started by Henry Clay Smith. This was the subject of a dissertation which he offered in 1939 at the John Hopkins University for the degree of Doctor of Philosophy. The research was carried out under the guidance of Sidney M. Newhall, and though the results were published as late as 1943, historically speaking, this research was conceived and executed much earlier.

In his experiments, Smith's purpose was to discover what relationship existed between colour discrimination and age, covering a wide range of ages. He was aware that spectrophotometric techniques would give him accurate results but he knew that these could only be used with well-trained subjects who were willing to spend a long time in the laboratory. He did not make use of colour blindness tests nor Holmgren wool discrimination tests, neither did he dabble with the naming of colours. He wanted to use a technique of matching where the least possible use of language was made and the whole test was arranged in such a way that subjects were capable of understanding the technique after very simple, brief, instructions. Lastly, he aimed at using material where quantification could be employed.

I. General - He selected every second member of a special Munsell series of a hundred hues which had been prepared with minimal differences in saturation and lightness throughout the colour gamut. This series of fifty papers formed his Hue series. A further five hues were chosen at even intervals around the gamut for his hue standards, (5.6 yellow, 4.0 green, 4.4 blue, 7.6 purple, 2.9 red). Likewise five standards and a series of twenty-seven papers were selected from Munsell stock which varied in saturation but not in hue or lightness - these formed his Saturation series which, incidentally, was of a red hue. Finally, an achromatic Lightness series of fifty papers and five standards were selected.

Each of these three series was tested as follows : The five standards were mounted on one drum and the complete series on another and the subject was required to decide whether the single exposed sample from the complete series matched the single exposed standard sample. The randomly arranged complete series was successively exposed for each of the five standards, and note was taken of any declarations of matching by the subject from which an error score was later calculated. In any experiment, of course, only one sample physically matched the standard.

199 subjects were tested chiefly from elementary and high schools but some came from John Hopkins University and others from aged men and women's homes in Baltimore. There is a predominance of subjects under 30 years of age as can be seen from the Table :

Age 5 - 9	43 subjects
10 - 14	34 subjects
15 - 19	39 subjects
20 - 29	32 subjects
30 - 64	19 subjects
65 plus	32 subjects
Total	199 subjects

Note that in the younger four groups there were approximately 35 persons per sub-group, while in the age groups with the larger age spans, there were only 19 subjects, and 3~~5~~2 subjects respectively.

The visual angle subtended by the samples varied from $3\frac{1}{2}$ to $7\frac{1}{2}$ degrees according to the inclination of the subject.

Illumination was that of natural daylight with an intensity of not less than 10 foot candles, and not more than 20 foot candles.

This level of illumination was found to be the best after experimenting with other higher levels.

The treatment of data was in terms of the measure of 'skewness' and 'kurtosis' with a 'standard error' for the total distribution of scores for all ages in three dimensions. The measures indicated that all three distributions had a marked tendency to be skewed and platokurtic. It was found that using the mean and standard deviation correlation techniques was highly questionable with such data, and therefore to supplement this, the data was also presented in graphic form.

The means of the composite scores showed a rapid improvement in every colour dimension (i.e. hue, saturation and brightness) from the youngest group (i.e. 5 years old) to a maximum of 20 to 30 years of age. An analysis of data given showed that in no dimension was there a significant difference in the 20 to 30 group or in the 30 to 64 group. Thus, from the data presented there was no reliable evidence for a difference in colour matching ability between the ages of 20 and 64, although the trend was in favour of the younger group. However, there was evidence that the ability to match colours improved from childhood to maturity, and continued unimpaired until 64, then declined in old age.

Furthermore, the youngest and the oldest groups had the largest standard deviations in the distribution of composite scores, and the middle age groups had the smallest.

The age trends for hue, saturation and lightness were almost identical. It appeared, therefore, from this research that any factors operating to cause age differences in colour matching, affect the three colour dimensions almost equally. However, when the pattern of scores for each individual dimension (taking for example discrimination in green, purple, blue or red) was examined, it was found that the shape of the curves were the same for all except the yellow curve. In the oldest group the score in yellow was markedly superior to the score in other hues and was also superior to the middle age group's yellow score. This superiority was markedly different from the trend in the other four hues and, as an explanation, Smith quoted the yellowing of the lens. The relative difficulty of the individual items was as follows :-

1. The easiest to discriminate as far as hue was concerned were firstly, yellow, then green, purple, red and finally blue.
2. The data appeared to show female superiority in colour matching in the younger (i.e. up to 14) age group.
3. On the other hand, in the older group there was male superiority in the scores for saturation and lightness.

II. <u>Comparison of the scores on the colour discrimination test with performance on Ishihara</u> - Smith gave the pseudo-isochromatic plates to every subject who took part in his colour matching test. To determine the relationship between the number of misreadings on the Ishihara test and the colour matching ability, the mean scores in hue, saturation and lightness were divided into three groups. Thus, he had subjects who made no mistakes on the Ishihara, those who made one, and those who made two or more. As in the case of sex differences, the data appeared to fall into two groups - those of under 15 and those over 15.

In the younger group, 5 out of 6 comparisons were in favour of those individuals with a <u>greater number</u> of mistakes on the Ishihara. Thus the trend, though not statistically significant, was towards a superior colour discrimination ability among those with two or more mistakes. In the older age group, on the other hand, 11 out of 12 comparisons were in favour of superior ability in the group with no mistakes on the Ishihara.

To explain these age differences between the Ishihara scores and their relationship with discrimination, Smith suggested that faulty number reading ability, inadequate comprehension or poor perception of the number configurations in general, rather than any physiological explanation, might account for the discrepancies. He supported this explanation by saying that a larger number in the youngest group than in any other age group made one mistake on the Ishihara.

Among his colour blind subjects, he found that there were wide individual differences in matching abilities. However, it appeared that the colour blind were as good as the normal subjects in discrimination. To explain this he advanced the hypothesis that, though the chromatic papers were of equal lightness for the normal subjects, they were not equal for colour blind, and thus these individuals could discriminate hues and saturations in terms of brightness differences.

III. Effect of Education and Training on Colour Discrimination - Smith also classified his results to test whether education and training are relevant factors in colour discrimination. Subjects were divided into two groups - those with low, and those with a higher educational achievement. He found that educational achievement was related to ability to match colours, and even went so far as to say that differences in intelligence might be factors operating to cause the difference in matching ability. Then the idea of training was tested. He compared the data on his tests with data for 5 employees of Munsell Color Company and found that in each case their scores (i.e. the employees) were much better than the average score, even for the top educational group. However, a comparison of the scores of a group of art students with other groups at the same age, gave no positive results.

IV. Conclusions - In his discussion, Smith compared his work with previous work and concluded that his results were consistent with those of Tucker (1911) who found that blue was the most difficult to match and yellow the easiest. He claimed that his work confirmed the work of Gilbert (1894) who showed that there is a steady improvement in saturation matching up to the age of 17, and confirmed also the conclusions of Collins (1925) and Hayes (1926) that colour blind subjects have a wide range of highly individual defects. Lastly, he concluded that his research also confirmed the work of Davidson and Pierce who revealed that there was a positive relationship between colour matching, and intelligence and experience.

He considered that his work was consistent with the general theory of Koffka who said that growth of colour discrimination started with the ability to differentiate light from dark in the newborn infant, then the discrimination of warm and cold colours developed, and finally there was differentiation into the primaries red, yellow, green and blue. Such colour discrimination was thought to be due to the relative influence of maturation on learning at the various stages of development. He thought that maturation of the pigmental retinal processes or the visual areas of the central nervous system might account for the decline in ability in old age. On the other hand, if maturation of the physiological apparatus for colour discrimination were complete at birth, the growth of discrimination ability might be due entirely to learning or at least to a more generalised maturation. However, maturation of visual processes is only a minor point and plays a small role in the account of age differences and colour discrimination after infancy. Therefore an explanation in terms of learning processes became necessary. He found discrepancies between his work and that of Houston (1932) who assumed that colour discrimination varied continously, whereas Smith's distributions showed that it is skewed and platokurtic.

In conclusion he made the following points :-

1. Ability to discriminate colours by matching improved rapidly from 6 years of age to adolescence, till at 25 it reached a maximum. In the middle age group, from 30 to 64 there was no significant decrease in this ability, but there was a definite decline in the very old.
2. Ability to match colours did not improve with any increase in illumination beyond 10 ft. candles.
3. Individuals with over average educational level for their age obtained significantly higher scores than those below, suggesting that high intelligence was related to improved colour discrimination.
4. There was a sex difference, since females between 5 and 11 were better than males at matching hues and saturation. Above 14 years, the males were superior to the females.
5. The correlation between mistakes made on the Ishihara test and colour matching ability was approximately zero, when the entire group was considered. However, in the younger people (up to 15) there was a positive relationship between the number of misreadings on the Ishihara and discrimination of colour.
6. It was suggested that age differences in colour matching ability were predominantly determined by attitudinal and learning factors rather than by receptor physiology.

(b) The Work of Joseph Tiffin and Hedwig Kuhn

I. General - This is the second important piece of research into the state of colour discrimination with age. Large numbers, running into thousands, were tested, thus overcoming the basic criticism made against Smith's work. Yet this research was never considered to be of any value by later critics. The authors of this new approach were penalised by the lack of a suitable description of the test they used. Tiffin had to pay quite dearly for this omission, as can be seen from the way that references to age and colour vision appeared and disappeared in the various editions of his text-book "Industrial Psychology". In its 1942 and 1947 editions, this topic was included. However, in subsequent editions there was no mention of these results at all.

Unlike the work of Smith, which went unnoticed, Tiffin's results and conclusions caused quite a stir in the United States. For example, 3 researchers were instigated to check on Tiffin's general conclusion at that time and each in turn managed to diminish it step by step till Chapanis' work finally reduced Tiffin's contribution to nothing, ascribing the results he obtained to the use of an invalid test. However, I consider this work to be very important indeed.

The paper 'Color Discrimination in Industry' was read in June, 1942 at the 93rd annual session of the American Medical Association's Opthalmology Section, and appeared in print in November of the same year in the Archives of Opthalmology. The authors assumed that colour discrimination and colour vision, as such, were very important for industry in general (and not only for operating signals or coloured lights). Their results were obtained by testing some 7,000 industrial employees.

A survey was conducted in a sheet and tin mill where several tests were administered, including simple four-item red-green and yellow-blue discrimination tests. The paper can be considered in four sections on the basis of his conclusions which, to say the least, were startling and unexpected.

II. Findings -

(i) Most items of the red-green discrimination test were passed by 55.3% of the 7,000 employees tested, but Tiffin and Kuhn emphasised that this did not imply that the 44,7% who failed were colour-blind. All that could be concluded was that such people had less ability to discriminate colours than the others. They also pointed out that their results were based primarily on a test of red-green discrimination. Later they added a four-item test of yellow-blue discrimination when testing about 500 employees but thought that this number was not sufficiently large for detailed statistical analysis. About 4½% of these 500 who missed no items on the red-green test, missed all the items on the yellow-blue test. From this fact they concluded that 4.4% of these 500 employees were more deficient in yellow-blue discrimination than in red-green. The suggestion was made, therefore, that, (contrary to general belief) the results showed that weakness in yellow-blue discrimination might in some cases be unique and that it might be present without a significant weakness in red-green discrimination.

(ii) Colour vision is often necessary for employees in jobs which have not been previously thought to require colour vision.

What follows refers only to the red-green discrimination method of scoring. The test or merit rating on the job was made by supervisors, and three grades of merit were given - very high, average, and low. The high group contained one third of employees who had been rated best in performance. The other two groups containing one third each of those who were rated average and low.

Employees were further sub-divided into 'five-year' groups. The age group continuum was from 20 to 70. Except for the two extreme groups (i.e. 20 and 70) neither of which contained many employees, at every age level, the group of employees who had been rated high in job performance, contained a significantly larger percentage of employees who passed the test of colour vision. Tiffin and Kuhn concluded that a fairly high level of colour discrimination was frequent among mill workers who have been rated high in job performance, than among employees who had not been so rated, in spite of the fact that not more than 5% of the jobs involved were considered to require colour vision. Two explanations were given :-

(1) It was possible that colour vision was required of many more employees than was thought. Every employee was also expected to be able to distinguish between company forms and charts which were often printed on different coloured paper for easy identification. Again, everyone in the mill was required to know

something of the colour code which identified the plumbing in his plant. For these reasons, Tiffin and Kuhn thought that the inability to handle the related aspects of the job properly accounted for the lower merit rating given to the employees with deficient colour vision.

(2) The second explanation they gave was that perhaps colour vision was a barometer of general health and physical condition to such an extent, that employees with reduced colour vision also suffered from general physical debility so that, in general, they were unable to perform their respective jobs satisfactorily - hence the low rating received from the supervisor and the poor scores in the test of colour vision. They quoted that excessive use of tobacco and alcohol reduced sensitivity to colour.

Whatever the final explanation, the statistical conclusion was inescapable, that the least desirable employees, from the point of view of general job performance, were those who failed the colour vision tests and it followed as a natural corollary that such undesirable employees could be identified by the colour vision test before they were put out of a job. Smith and Kuhn then claimed quite strongly that industry had in such a colour vision test, a reasonably satisfactory device for predicting whether a man hired for general mill work would be successful on the job.

Another approach was made by studying the percentage of employees on each of seven specific jobs who were able to pass the test of colour vision. For example, the percentages for foremen and for clerks showed a statistically significant difference from the plant average, and these differences could only be explained by random or choice variation. They concluded that the results indicated a real individual difference between the employees in general and the specific employees. Tiffin thought that it was understandable that the foreman group should contain a larger percentage of men with good colour vision, since these men had been familiar with all kinds of details pertaining to the jobs, among them the colour codes of the plumbing system, colour lights and signals and coloured charts. It was, however, more difficult to explain why clerks were superior in this respect, particularly in view of the fact that they were not on ly better than the average mill employees, but were clearly better than the foremen as well. Tiffin was unable to explain this satisfactorily.

(iii) Colour vision shows a marked reduction with increasing age, making it necessary for an industry to be certain that persons to be trained for certain jobs should have more adequate ability when the training begins.

The analysis of age effect on ability to pass the colour vision test showed the following percentages for the various age groups :-

Age Group	Percentage of Pass
20 - 25	74
26 - 30	70
31 - 35	68
36 - 40	58
41. - 45	42
46 - 50	39
51 - 55	28
56 - 60	20
61 - 65	28

Thus, though for example 74 per cent of employees in the 20 - 25 age group passed the test, only about 30 per cent passed it at the age of 55 and over, representing a ratio of more than 2 to 1 for those employees at the age limits. Tiffin and Kuhn concluded that colour vision must be a function which was lost in later years and also thought that the less that there was at an early age, the more quickly the loss with age would reduce functional efficiency on any job that required this particular skill.

They then added that there were many industrial jobs in which this loss could be serious indeed, quoting operatives in the modern colour printing press, and went on to say that the fact that there is a loss of colour vision with age should be noted by every personnel manager, who should make certain that every employee who received special training had sufficiently good colour vision at the outset, so that even after some of his ability had deteriorated, it would remain good enough for satisfactory job performance.

(iv) Colour vision exists in greater or lesser degrees and is not an all-or-nothing proposition of the type assumed by the classification of persons into colour blind and normals. - Indeed the great majority of persons were thought to be somewhere between the two extremes. The authors talked about a continuum and said that in their survey no employee tested was completely colour blind, as this was very rare, probably as rare as the extremely high degree of colour sensitivity possessed by the few gifted painters or by others whose work is dependent on their possessing a high level of colour discrimination. His results implied that a test was required for industry which could measure the degree of colour acuity rather than the usual dichotomy of "normal" or "colour-blind". It was not sufficient to know that a trainee did not belong to the colour defective class.

2 : 3 STUDIES BASED ON PSEUDO-ISOCHROMATIC TESTS ALONE

The works of Marie Lou Boice (1948), Kleemeier (1952) and to a lesser extent, that of Chapanis (1948, 1950) were the first researches in which tests for inborn colour vision anomalies were used exclusively for detecting changes in colour discrimination due to age differences.

The researches of Boice and of Kleemeier were undertaken with the express purpose of assessing the validity of Tiffin's and Kuhn's results, and to examine some aspects of their conclusions. Indirectly they also discussed Smith's contribution.

It is rather surprising to note that when these two writers discussed the results obtained from pseudo-isochromatic tests, our attention was not drawn to the fact that the colour discrimination tests used by Smith and Tiffin were rather different from tests used for the detection of colour vision defects - in terms of construction, design, task involvement, and the amount of colour discrimination that was required for high scores on such tests.

Chapanis (1948), who was interested in the problem of acuity in addition to colour, gave his first conclusions on this topic as incidental findings, claiming that his results showed that neither visual acuity nor age were related to performance on five colour vision tests.

In his 1950 paper, Chapanis stated that when he was invited to prepare an exhibit for the Baltimore Sesquicentennial Exhibition, it seemed an unusual opportunity to collect extensive data on the controversial problem of the influence of age upon acuity and colour vision.

(a) Study of Boice, Tinker and Paterson

The paper 'Color Vision and Age', appeared on the 30th June, 1948 in the American Journal of Psychology. It was a summary of a thesis presented by Marie Lou Boice at Minnesota University.

I. General - The Ishihara 32 plates and the American Optical Co. 42 plates were used. Subjects were members of the Faculty of the University of Minnesota and were a highly selected and homogeneous group in respect of intelligence, formal education, occupation and socio-economic status. All tested were of academic rank - in total 236 men.

The distribution was as follows :-

	Age Group	Subjects
1.	20 - 29	80
2.	30 - 39	78
3.	40 - 49	29
4.	50 - 59	28
5.	60 +	21

Constant illumination was provided by using a 100 watt GENERAL ELECTRIC daylight bulb. The standards for diagnosing colour blindness were the same as those of Thomas and Foster. On this classification, 7.2% subjects of the total population were found to be colour blind.

II. Age and Colour Efficiency - The percentage distribution of colour blind by age decades showed an irregular but small proportion diagnosed from the age of 20 to 59. The authors interpreted the irregularities as being due to experimental error. However, almost one quarter of those aged 60 were found to be 'colour blind' even though the sample contained only 21 subjects. To check this possibility an additional 19 men aged 60 or over were tested, and this sample likewise produced an unusually high percentage of colour blind men (i.e. 15%) or combining both samples into a group of 40 subjects, the percentage produced was 20.

III. Interpretation of Results - Boice et. al. criticised Tiffin for saying in the 2nd edition of his book 'Industrial Psychology' that both his studies agreed that colour vision deteriorated with age, (that is from his work with the Orthorator (1949) and that on the Telebinocular). They also claimed that other inferences made by Tiffin, taken in conjunction with their results, left his conclusions in an uncertain and unsatisfactory state. What was necessary was a large-scale research using both men and women where intelligence, education and occupation levels could be controlled and where adequate colour perception tests such as the Ishihara could be used. Suggestions were given for a longitudinal approach (i.e. testing people at 45 years, and re-testing them at 5 or 10 year intervals over a period of 30 years).

(b) Summary of Paper by Alphonse Chapanis entitled 'Relationships between Age, Visual Acuity and Colour Vision' which appeared in 'The Human Biology', February, 1950.

This study was concerned with the relationship between age, visual acuity and colour vision in an unselected sample of individuals. Chapanis was familiar with the previous studies on Age and Colour Vision, and quoted the results of H. Rydin (1927), Ferrara (1940) and others who attributed the ageing effect on colour discrimination to lens changes. He was aware of the work of A.I. Collins (1918) and L. Cronstedt (1937) who found that even among 'harmless colour blind' people, presbyopia, hyperopia and hyperopic astigmatism were prevalent.

As we are not directly interested here in the relationship between visual acuity and age, only his results on Age and Colour Vision, and acuity and colour vision, will be considered.

I. Testing Situation - Data was collected at the Baltimore Sesquicentennial Exhibition in 1947. The subjects were a cross-section of typical exhibition visitors with a variety of educational backgrounds, intelligence and occupations. Altogether there were 574 subjects, 302 males and 226 females between the ages of 7 and 77.

All tests were illuminated by a Macbeth daylight lamp 5000° K. at the test surface giving an illumination of 30 ft. candles. Tests used were as follows :-

1. An experimental 6-plate pseudo-isochromatic test; devised by the author;
2. C.J. Bostrom and I. Kugelberg Test 1944;
3. Ishihara test for colour blindess 9th edition;
4. Dvorine's colour discrimination screening test 1947;
5. Pseudo-isochromatic plates for testing colour perception printed by Beck Engraving Co. 1940.

II. Results - Subjects are scroed according to the number of correct plates, the highest possible score being 85. Chapanis found that there was a tapering-off at the 36 to 40 score, and below that again there was a build-up of subjects. The data pointed to two distinct distributions, one with scores above 40, and one with scores below 36. As there were only around 46 of the 574 individuals tested in this low score group, (i. e. 36 - 40), Chapanis concluded that all those who read 30 or fewer plates correctly were colour blind.

Following a rather more rigorous classification these 46 individuals were probably true dichromats. Approximately 13.2% of all males and 1.66% of all females fall into this category. This, of course, was a considerably higher proportion of colour defective individuals than was normally found in a randomly selected population. To explain this large number, he suggested that, as many of these examinees had long suspected their colour deficiencies, they had decided to take these tests in order to find out their defects. Negroes also earned somewhat poorer scores than whites on this test battery.

III. The Relationship between Age and Colour Vision - Chapanis correlated the total number of plates read correctly on all colour vision tests with age, and found that the linear correlation was equal to + O.076 and the curvilinear correlation was equivalent to 0.177. Neither correlation was significantly different from 'Zero'. Of course, this finding was suspect because of the very obvious bimodality of the test data; consequently he discarded all those with a score of 50 or more and found that the correlation between age and colour vision was still low ($r = 0.15$). However, by means of the 'F' test, it was found that a curvilinear relationship was more consistent with the data than a linear relationship and that the curvilinear correlation was significantly different from zero. But if the data of children of 14 years or under were omitted, the two correlations again were insignificant. Hence it appeared that any correlation between age and colour vision that might exist was due to the young examinees who do not perform as well as adults on these tests.

IV. Correlation between Age and Colour Vision Scores for Individual Tests - Because there was some likelihood that individual colour vision tests might have given different results, correlations were computed between age and scores on single tests. Only one of them was significantly different from 'Zero'. The scores on the individual colour vision tests were all bi-modal in essentially the same manner as the scores for all the tests together. He concluded that the analysis by individual test did not alter his conclusion on the association between Age and Colour vision.

Chapanis found no significant relationship between visual acuity and colour vision test performance. On the relationship between Snellen rating of acuity and colour vision scores he said that if the markedly colour deviant individuals were excluded, there did not seem to be any positive association between visual acuity and performance on these tests. People with better acuity tended to make rather fewer errors. He concluded that the relationship between acuity and colour vision is statistically significant but so low as to be almost insignificant.

(c) Study of R. W. Kleemeier

The last work to be discussed in this section has only an indirect bearing on this research. This was the study which R. W. Kleemeier published in the Journal of Applied Psychology in 1952 under the title 'The Relationship between Ortho-Rater Tests of Acuity and Colour Vision in a Senescent Group'.

In this paper the author attempted to explain another of the startling findings of Tiffin and others. In 1950, Ely, Kephart and Tiffin published results of a survey under the title 'Ortho-Rater Norms and Sex Differences'. In this survey, of the 7,500 male and 2,500 female industrial employees who were tested, an unexpected difference in colour vision scores in favour of the men was found. This difference was significant at the 1% level. The authors were aware that this finding was contrary to long accepted theories and facts about the distribution of colour blindess among the sexes.

In his discussion of his paper, Kleemeier thought he had an explanation for these findings. The poorer visual performance of women in the industrial group on the colour vision tests was obvious, and he ascribed it to acuity problems. Thus, it was not colour perception but rather deficiency in visual acuity that was thought to be the cause. The women simply could not see the colour chart as well as the men. He also referred to Tiffin's article in 'Industrial Psychology' (second edition 1947) on the dimunition of colour sensitivity with age. He thought that, in view of the possible confusion of these results with uncontrolled visual acuity, the reported age trends in colour vision were open to question. It would appear that any attempt to ascertain the relationship between colour vision and age would be successful only if visual acuity were somehow to be controlled. This was particularly true when pseudo-isochromatic colour tests, such as the Ishihara or the Ortho-Rater were used.

2 : 4 STUDIES INVOLVING METAMERIC MATCHINGS

The study of colour discrimination using metameric matching is one of the more sensitive methods of measuring the differences in individual sensory discrimination. The essential idea underlying such matching is that two physically different stimuli, as far as spectral composition is concerned, may produce an identical sensation in an human observer.

Metameric pairs can be lights or surface colours. Of the following researches, two employed lights, and the third study involved object metamerism. The phenomenon of metamerism will not be described any further at present, as it seems, from the accounts given in the papers to be discussed, that the writers did not look upon their work as representing research into age and colour vision by means of metameric pairs. Of course, this lack of understanding prevented them from obtaining any additional results from their data relevant to the problem of age and colour vision.

R. Brown used the anomaloscope mainly for the detection of major colour defectives, and only incidentally for the measurement of minor differences in colour perception.

Though F. L. Warburton used a dochroic test, his interest lay mainly in the practical aspects of colour changes due to age.

B. Boles-Carenini looked upon the Nagel anomaloscope as a device in which the 'form' element present in the pseudo-isochromatic plates was excluded.

(a) "Investigation into the Colour Vision of School Children" - an Ed. B. thesis by Robert Brown.

A summary of this thesis appeared in the Br. J. Ed. Psych. in 1950. The aim of the investigation was to find out if the incidence of colour vision defects and the pattern of normal colour discrimination was the same in children as in adults.

Brown used the Ishihara Plates and Pickford's colorimeter. 525 boys and 252 girls between 5 and 15 were tested. 84 boys and 36 girls were tested both on the Ishihara and on the anomaloscope.

He found that the frequency of colour defectives among West of Scotland school children was similar to that found by other investigators and followed the European pattern - 7.43% for boys and 0.45% for girls. This showed that the distribution of colour vision defects followed the same pattern in children as in adults.

A further investigation carried out along similar lines showed that there was no significant difference in the colour vision of school children and adults. He used the Chi-squared technique to compare the results of the red-green test and the yellow-blue test on boys, with similar results for normal men, and finally girls with normal women. He found that the probability score was 0.1 which was not significant. No actual score for the adult population was given, nor was age stated.

(b) The Study of F. L. Warburton into 'Variations in normal colour vision in Relation to practical matching'

Warburton's was a practical problem. He accepted the advice of the Colour Group in their Report on 'Defective Colour Vision in Industry' (1953), that a small proportion of personnel in industry must have exceptionally good colour discrimination. It was considered inadequate merely to eliminate those with defective colour vision. Variations in normal colour vision also had to be taken into account.

The most important of these variations was attributed to absorption by the macular pigment in the eye. Warburton then quoted Wright's (1946) research, where chromaticity points for a 'white light' were measured for 36 observers, and it was found that such points lay in an elongated region, starting about the actual position in the C. I. E. diagram for illuminant 'B', and moving towards the yellow spectrum locus. Warburton thought that this work by Wright suggested that the influence of degrees of macular pigmentation in practical colour matching was considerable. He gave as an example the possibility that the colour temperature variations for illuminant 'B' in this group might be between 3,500 and 7,500^{o} K. for the individual observer.

I. Testing Situation and Procedure - 250 observers, mainly connected with wool and related industries, were examined - all were people who came to an exhibition on 'Light and Lighting'.

The test used was that devised by the Bradford Dyers' Association. It consisted of a slightly dichroic standard pattern which was to be matched against one of a numbered range of dichroic patterns dyed with dyes different from those used for the standard; the position of the match varying with the observer and the illumination under which it was viewed. The patterns in the numbered range were each about $2\frac{1}{2}$ ins. long by 1 in. wide and varied in hue from a dull green (No. 1) to a reddish brown (No. 13). They all appeared redder in artificial light than in daylight. This set of patterns was made available to the author and spectrophotometric data obtained using the Hardy Recording Spectrophotometer. The patterns were illuminated using illuminant 'B' and visitors asked to fill in on a card the number of patterns in the range which matched the standard. The card also contained spaces for additional information such as the age and occupation of the observer.

No preliminary test for defective colour vision was made, but it was thought unlikely that any such person would be successfully engaged in work closely related to dyeing, at any rate without being aware of his defect. Comments on the difference between the standard and the nearest pattern in the range also indicated the more serious cases of anomalous trichromats who could then be eliminated, since from the nature of the test it was unlikely that they would find a satisfactory match in any of the range of patterns. A few cards were returned by visitors who were known to have defective colour vision. The differences between their cards was remarkable.

II. Results - Altogether 241 observers participated, the age distribution being as follows :-

Age Group	Subjects
16 - 25	32
26 - 35	55
36 - 45	86
46 - 55	51
56 - 65	17

Since for technical reasons, it was impossible to determine the match which would be found by a standard observer, the effect of the change of illumination on the position of the match was determined using the author as a standard. The patterns were observed under illuminants 'A', 'B' and also using an under-run Macbeth colour matching lamp to give an illuminant close to 'C'.

These were the matches :-

Illuminant	No. of sample
S A	3
S B	8
S C	12

The results for the total population were as follows :-

The mean position was midway between the points representing the chromaticities of patterns 8 and 9. For practical reasons then, he accepted the mean as the width represented by patterns 8 and 9. Total matchings ranged from patterns No. 5 to No. 11, that is, equivalent to a change in illuminant from halfway between S A and S B to, at the other extreme S C, and thus similar to the colour temperature variations expected from Wright's work.

The age variation on the other hand showed the following distribution :-

Age Group	Mean Pattern	Range	Equivalent Illuminant
56 - 65	7 - 8	5 - $9\frac{1}{2}$	SA to SB
46 - 55	8	5 - 11	
36 - 45	8.5	5.5 - 11	
26 - 35	9	6 - 11	
16 - 25	9.5	8 - 11	SB to SC

A change in position of the mean was definitely significant for alternate age groups but not always so for adjacent groups. This analysis by age groups provided concrete evidence that there was a definite increase in the degree of pigmentation with age.

An attempt was also made to see whether there was any appreciable difference in the distribution of pigmentation between dyers and the general public. The view has occasionally been expressed that when a young dyer is appointed, he is often selected because his colour vision is similar to that of the senior dyer. The latter often has a large amount of pigmentation with the result that when the second man in his turn appoints an assistant a still more pigmented observer is chosen. This has been tested and no significant difference from the general survey found, except that there were relatively few cases of extreme over- or under-pigmentation. This would imply that the biased selection of dyers suggested above was not a frequent occurrence, partly because many senior dyers, particularly those working with large firms were aware of the change that came with age. Supporting evidence for this unbiased selection was found from the fact that the whole range of pigmentation as found among dyers occurred among the personnel of one firm.

III. Interpretation and conclusions - Comparison of results showed that influence on practical colour matching of the variations in retinal pigmentation found by Wright equalled the change in light source required to produce a chromaticity difference on the W.D.W. System as shown for different observers for the 'National Physics Laboratory White'. The difference was seen to be considerable and this standardisation of the colour of light sources for colour matching is of very little value unless the observer is also standardised somewhere near the average pigmentation level. A further deduction was that less stringent limits were required for the colour temperature of a source having a 'black body colour' than were required for a source which departs from a black body energy distribution.

The analysis also showed that, as a long term policy, such standardisation of the observer would not give rise to any serious difficulties since 60% found matches ranging between patterns 8 and 9. Immediate standardisation would, of course, cause hardship to existing personnel who did not comply with the standard. The demonstration of pigmentary variations raised the question of an appropriate test for selecting colour workers. The most satisfactory absolute test would be some form of 'vecticolorimeter' which would enable the subjects' white point to be plotted on the chromaticity chart. Such an instrument, however, would require to be used under proper laboratory conditions. For general use, an approved form of the test used in the present investigations would appear to be suitable.

(c) B. Boles-Carenini

The last research, involving metamerism, is the study of B. Boles-Carenini. The account appeared in an Italian Journal of Opthalmology (1954) under the title 'Del Comportamento del Senso Chromatico in Relzione All'ete' (Trans. 'On the Behaviour of the Chromatic Sense in Relation to Age').

Boles-Carenini gave an extensive summary of previous studies, and stated that his aim was to examine results reached by the various workers, using apparatus in which the 'Form' element present in the pseudo-isochromatic plates was excluded.

I. Method of Study - The Nagel anomaloscope was used, which eliminated the interpretive factor of form, and at the same time enabled a statistical analysis of results to be made which could furnish an exact picture of the variations possible in recognition and in colour discrimination.

170 subjects, free from ocular defects, and possessing normal visual acuity were tested. For every subject, ten observations were made in the red-green equation. The means and standard deviations for the various age groups obtained from the Nagel instrument were as follows :-

II. Findings -

	Age Group	Number	Mean Setting	Standard Deviation
1.	20 - 29	30	40.1	2.38
2.	30 - 39	30	39.9	3.43
3.	40 - 49	30	39.7	3.09
4.	50 - 59	30	39.9	3.60
5.	60 - 69	20	38.2	3.00
6.	70 - 79	30	34.4	3.43

Two statistical analyses were attempted, one for the measure of dispersion, the other for central tendency.

For analysing Dispersion, Boles-Carenini used Fischer's 'Z' formula where $Z = 2.3025 (\log \sigma_1 - \log \sigma_2)$ and found that significant differences were obtained (at the 1% level only) between the 20 - 29 age group and all other age groups. Differences in S. D. between other age groups were not significant when compared with each other.

The results from the analysis of Dispersion prompted him to say that the data obtained can be thought of as coming from two distinct sources :-

a) from people below 29 years of age; and

b) from those over 29years.

He then attempted to analyse Central Tendency. The means for the first 4 age groups were constant, but in the last two there was a shift to the green end of the equation. To measure the significance of this shift he compared adjacent age-groups using Student 't' scores and found significant differences only between the fifty and sixty year olds and between the sixty and seventy year olds. Thus he concluded that the first four means came from a homogeneous population but that the last two age groups did not.

III. Conclusions - A difference existed in the distribution of the values of the Rayleigh equation. In biological terms, this implied that among people under 30 years of age, there was a probability of finding subjects nearer the norm in red-green discrimination. On the other hand, in age groups over 30, anomalies were more likely to be found in red-green discrimination.

If, however, comparison of the means was considered, it was emphasised that statistically appreciable differences were in evidence only after 60 years of age, and took the form of a significant decrease in the perception of green.

Boles-Carenini then added that the possible pathogenic mechanism of such an anomaly might be attributable to senile sclerosis, which was more marked after the age of 60.

His final conclusions were that the statistical analysis of results proved that no relationship existed between age and chromatic sense in subjects over 20 and under 60 years of age. Thus age had no influence on the ability to recognise and discriminate between colours up to the age of 60, after which the possibilities by dyschromatopsias occurring would increase even in subjects previously possessing normal colour vision. Such conclusions were confirmation of the work of Boice, Tinker and Paterson.

2 : 5 SUMMARY OF WORK DONE BY K. JANOUSKOVA PUBLISHED IN CZECH JOURNAL OF OPHTHALMOLOGY, 1955, UNDER THE TITLE 'BAREUNE VIDENI A VEK' ('COLOUR VISION AND AGE')

The study began in Prague when the author systematically examined colour vision differences to find out if there were any differences in colour vision in the various parts of the country. Only after testing around 7,000 patients did she become aware that the differences between children's and old people's scores were greater than variations from district to district.

This prompted Janouskova to seek answers to three fundamental questions :-

(i) Was everyone born with the same ability or inability to distinguish colour ?

(ii) Was there a loss of colour perception in the old ?

(iii) Lastly, if there was a loss, how much was it due to lens deterioration ?

(a) Growth of Discrimination in Children

I. General - To find the answer to the first question the author tested 4,450 children between the ages of 3 and 12 years of whom 2,300 were girls and 2,150 boys. All these children had normal fundi and only very small refractive errors. The results were classified in 8 age groups. There were very few of the children in the 3, 4 and 5 years' group and to supplement this, a study of 30 additional kindergarten children was undertaken. No babies were included (i.e. from 60 to 150 days).

Each child was tested in two ways :-

(i) by employing a naming test, which consisted of embroidery silks in seven fundamental colours, and

(ii) by using the American Optical Plates (1942) and the Ishihara Test to test what the author calls 'differentiation'.

To find out whether children were ready to do this second type of test, a card composed of grey, white and black dots was employed, and the children asked to trace the path. In this way any distortion of results due to the child's failing to understand the test was avoided. For children who did not recognise the figure, the test was still further adapted. A toy car was set before the child and the figures indicated. He was told that this was the road and asked to go along it with the little car. Alternatively, the child might be given a paint brush and told to sweep the path with it. Each child was examined separately. Often it was necessary to gain confidence first, but once this was achieved, there was no difficulty in testing.

II. Results obtained -

(a) Boys :

Among the three year olds, it was found that only 25% of the boys named the colours correctly, but 50% differentiated the colours correctly. There was an increase of correct naming and differentiating of colours between the ages of 6 and 7, where a great improvement was shown. This was attributed to the children being of school age.

From the age of 8 onwards, only 7 to 9% were found with faulty colour discriminations, and in the remaining older age groups, the data conformed to the percentages of defective colour vision found among adults.

(b) Girls :

Among the girls a similar rise in the percentage of correct responses from 3 to 6 years was found.

From 7 years on there was a great improvement which persisted till 12 years of age and corresponded to the statistical figures found for adults.

Up to 8 years of age no difference in colour perception due to sex was found, but on the whole, it seemed possible that, in girls, colour perception developed more quickly and became more acute.

As a general explanation for her findings, Janouskova suggested that children must be born with the basic physiological colour mechanism, but that maturation processes were at work, as evidenced from the progressive development of colour discrimination.

(b) Loss of Colour Perception in the Old

The writer was acquainted with the work of Tiffin, and to overcome Boice's criticism of the effect of such factors as intelligence, education and profession, she selected from 12,000 individuals, 1,478 men and 1,340 women, aged between 50 and 85, who satisfied the criteria mentioned above. The group was approximately equal in age but had the lowest attainment in education and intelligence. Colour vision was tested by means of the A. O. test (1942) in daylight and, where necessary, optical corrective glasses were used. No classification into deutans, protans, etc. was attempted as the test did not allow for this.

I. Analysis of Results - The incidence of those with 'defective' colour vision was as follows :-

Age Group	Male	Female
51 - 55	8.7%	4.5%
56 - 60	11.0%	4.2%
61 - 65	13.0%	6.5%
66 - 70	16.0%	7.5%
71 - 75	18.5%	7.0%
76 - 80	25.0%	14.0%
81 - 85	26.0%	12.0%

Thus, in the male population in the first age group only 8.7% had 'defective' vision, while at 85 the incidence was 26%. Of the women tested, 4.5% of the first age group were colour defective, and by the age of 80 this percentage rose to 14. There were approximately 150 to 250 people in each of the age groups.

In the study of the fundi of these subjects, it was found that among the 75 year olds in each group, approximately 60% had varying degrees of incipient cataracts or vaso-sclerotic changes in the retina. In the last age group (i.e. from 75 to 85), 75% had these defects. The results, then, showed a real deterioration in colour perception after the age of 60 in both men and women.

It was also noted that among people with higher intelligence the deterioration of colour discrimination with advancing years was rarer. In testing a group of selected people, with high academic qualifications, between 60 and 75, whose eye tests gave normal results the colour losses were found to be 18%, whereas, in the former groups, it was as much as 25%.

(c) Influence of Lens Deterioration

After first ascertaining that there was a definite dependence of the deterioration in colour vision on advancing age, the next point was to find out how much of this was due to lens deterioration. The writer was aware of Hess' (1908, 1909) and Ahlenstiel's (1943) work where a deterioration of colour vision due to lens absorption was shown. Furthermore, it was argued that the yellowing lens acted as a yellow filter in the eye and thus the colour sense deteriorated. On the basis of Tendelenburg's work the author conceived an idea of measuring this by her own test. She chose from among the 125 Ostwald colours suitable for this, shades of yellow, red, violet, blue, green and grey. For each standard, comparable colours of suitable hue were prepared, and for the grey tests a series of desaturated colours were used. Thus, in such a test, recognition by form (i. e. letter and numbers or by naming of colours) was eliminated and the recording of results made easy. The test was then presented to a 30 year old artist, to protanopes and to other known major colour vision defectives. Finally, it was presented to people in the 60 to 70 age group.

The qualitative changes in colour perceptions were evident. For example, a particular patient with incipient cataract saw yellows as orange-yellows to pale yellows, distinguished the bright reds correctly, but saw greenish yellows as green with no yellow. Green-blue and blue-green were seen as green, while violet-blues were seen as grey and blue-reds as yellow-reds. It seemed rather strange that patients with a cataract saw yellow-green shades as green and deep yellows as orange shades. This was explained by the adaptation of the eye to yellow colour. The patient was so adapted to yellow by the yellowing lens that only strong yellow, golden yellow and orange pigment appeared to be yellow.

With advancing age, still greater deterioration in colour perception was found, which could not be explained absolutely by the absorption of light through the lens. To examine this point further, a group of 180 subjects of 60 - 70 years with incipient cataract in one eye only was tested. The other eye was aphakic. The aphakic eye was found to distinguish violet and blue colours significantly better than the non-aphakic eye.

Thus it was shown that a clouded lens could alter colour perception and cause deterioration in certain parts of the spectrum. Again, not all deterioration in colour vision in aged people could be explained by dense lenses, and it was suspected that deterioration of colour vision in old age could be caused by changes in the receptor system of the retina or by changes in the structure of the cortex.

(d) General Conclusions

In any occupation where colour vision is an essential factor for the successful execution of work, it is important that there should be frequent testing, and this was advised for all such work.

2 : 6 STUDIES USING THE COLOR APTITUDE TEST (CAT)

Another test used by those doing research into the problem of age and colour vision was the Inter-Society Color Council Color Aptitude Test (which will now be referred to as CAT). This test was designed to measure colour skills acquired through training in making various kinds of colour judgments. In it, a person was required to discriminate small differences in saturation and this ability depended on experience as well as the inborn capacity to make such discriminations.

The CAT consists of four widely-differing hues, each represented by twelve samples differing in saturation, making a total set of forty-eight samples. Two such sets exist - one is mounted in four rows according to hue on a test panel, the other is packed in random order in a dispenser. The testee is instructed to find the colour match on the panel, comparing each loose chip as he removes it from the chip dispenser and the test is scored in terms of the accuracy of the saturation matches. The greater the accuracy, the higher the score. It is important to note that the CAT scores do not measure defective colour vision, since some colour defectives become highly skilled in making fine colour discrimination in certain hue regions and occasionally accumulate higher overall scores on this test than others with normal colour vision.

(a) The first researcher to use this type of test was Louise A. Quellette who presented her work for an M. A. Degree thesis under the title 'Age Differences in Colour Discrimination' at Fordham University, New York in 1955.

She tested 28 subjects, half of them male and half female, in the age group of 20 to 30 years, and another 28 subjects, 14 male and 14 female, between 75 and 85 years of age. The young subjects were students at Fordham University and their friends, while the old subjects were residents of old peoples' homes in New York. These subjects were private patients from a rather better than average socio-economic class. They were selected because of their age, general alertness, visual acuity and interest in the experiment.

The resulsts were analysed to determine age, sex and colour differences in discrimination. It was found that -

(1) Young people were better able to discriminate colours than the older people. There was a significant difference (beyond the .01 level) between the mean scores for the young group and the older group on blue, red, green and yellow.

(2) There was no sex difference in the ability to discriminate colour. Males and females were equally able to match colours.

(3) No one colour showed any decrease in discrimination more rapidly than any other. The older group found each colour more difficult to match than did the young group.

(4) For the entire group of subjects it was found that they did significantly better (at .01 level) on red than they did on green, and significantly better (.01 level) on yellow than on green. Green appeared to be the most difficult colour to match accurately.

(5) There were wide individual variations in colour matching ability as tested by this fine measure of colour sensitivity.

(b) In 1957 Jeanne G. Gilbert published a paper entitled 'Age Changes and Colour Matching, using the CAT'

I. Procedure - In her study she tested 355 unselected subjects of whom 160 were males and 165 females. The age range was between 10 and 93. Apart from Quellette's 28 institutionalised subjects and a few older men from the sailors' institute for old men, all old subjects were living in their own homes in the community. All took the test voluntarily and co-operated well. No differences in motivation at the various age levels were noted.

Most of the subjects were given the test under daylight conditions, but at few took the test under a Macbeth 100 watt lamp.

Starting at the age of 10, subjects were grouped according to sex and age decades, with the exception of the highest age group which included individuals both in the 80's and 90's.

The rated scoring system of Dimmick and Foss was not used, but raw scores were converted into percentages and thus presented for the sake of uniformity.

In order to determine what proportion of the difference between the groups was due to age, sex and colour, an analysis of Variance was performed. This indicated that the colours showed score differences which were reliable at the .01 level of confidence. Sex differences on the whole were reliable at only 0.05 level of confidence, and within the age groups there were also significantly reliable differences (.01 level) between performances on the separate colours.

Colour scores for sex and age groups indicated that all scores, with the exception of yellow (which showed a slight rise in the 40's), tended to show an initial rise from the 10 to 19 decade reaching a peak in the 20th year, and then showed a fairly steady decline up to the oldest age group. Yellow and red discrimination was superior to green and blue for all age groups. The curves showed a decline similar in form to those found in most other ability studies.

II. <u>Reliability of Sex Differences</u> - Females were found to be superior to males in total colour scores at the .05 level of confidence. However, females did not score significantly higher than males on all colours nor at all decades of life. No one decade showed significant sex differences in any one colour, although the total female group was superior to the total male group at .01 level of confidence in their ability to differentiate shades of red, and at the .05 level of confidence in their ability to differentiate shades of green. Females also attained higher total colour scores than males at .05 lebel of confidence in the 10 to 19 decade, at .02 level of confidence in 40th decade, and at .01 level of confidence in 70th decade. No other significant sex differences were found.

Gilbert said that no explanation could be offered for the superiority of the female sex at some levels but not at others. It was considered possible that their superiority in matching reds and greens might be due to the greater incidence of the undetected red-green colour weakness among males.

III. <u>Differences between the scores at the various age levels</u> -

(i) The 10th decade was inferior to the 20th decade on all scores and there was no significantly reliable superior scores over the other age groups until we come to the 50th decade where its superiority was found to be significant at the .01 level of confidence, and this remained so throughout the remaining decades. These results were consistent with the findings of Chapanis, Cooke and Smith that, up to 15 years children do not perform as well as adults in colour tests, although the reason for this was not clear. It seemed likely that it might be due to a developmental factor.

(ii) The 20th decade showed total colour matching scores which were significantly better than those of all decades with the exception of the 30's. As for the separate colours, the 20 to 25 year olds were superior at .05 level of confidence only in differentiating shades of blue, but were superior to all other decades at .01 level in their ability to differentiate shades of this colour. The results indicated that the ability to match colours developed in childhood, reached a peak in the 20's, and from then on showed a gradual and steady decline to old age. Although the decrease was steady, it was gradual, so that differences between adjacent decades were generally not statistically reliable. It was only when alternate decades were compared, that the results assumed statistical reliability.

Gilbert thought that her findings were not consistent with Quellette's (that no one colour decreases in efficiency more rapidly than any other). They were, however, consistent with the findings reported by Chapanis that there was a more rapid decline in ability to discriminate blue. Ability to discriminate shades of green likewise showed a relatively rapid decline with age, whereas the decline in ability to discriminate yellow and red seemed to be less rapid.

It appeared reasonable to attribute these losses to the filtering effect of the progressive yellowing of the lens with advancing age. The practical implications of these findings for colour workers in the industrial field were mentioned, and Gilbert supported Tiffin's suggestion that because of this ageing factor, it would be desirable for workers in the colour field to have an initial excess of ability to discriminate colours. She also considered it desirable that, in addition to the initial selection of colour workers, a periodic check should be made on the efficiency of workers in this field, particularly on those who worked with blues and greens.

2 : 7 CRITICAL APPRAISAL AND SUMMARY

At this point an evaluation of the contribution made by the research workers discussed so far is necessary. This will take the form of a rather lengthy criticism of the contributions of two pioneers in this field of study, and reasons will be given for the negative results obtained from studies employing pseudo-isochromatic plates alone. Finally, the remaining contributions will be evaluated and followed by a summary and general conclusion.

(a) Assessment of the contribution of H. C. Smith

This is rather a remarkable study as it departs from the previous unsystematic type of research carried out on colour discrimination. It is interesting to note that Smith was attempting to devise a new test, a forerunner of the colour discrimination tests that were to be used in the '50s and are so popular to-day, such as the Farnworth-Munsell 100 Hue Test. The methods of applying the test, and the care which was taken to establish the optimal illumination are to be commended.

However, in terms of methodology Smith's work suffers greatly. This is primarily because he did not take sufficient care to have equal numbers in his age groups. It is almost incredible that he took the trouble to test 30 - 40 subjects for each 5 years up to 30, and then for the next 6 sub-groups each covering a 5 years' span, he tested only 19 people. His interest was only in the growth and decline of colour discrimination, but he talked about colour discrimination through all the ages in his conclusions and even went so far as to say that between 30 and 64 there was no loss whatsoever.

Nonetheless, certain factors should be noted which have a bearing on the evaluation at the end of this thesis. These are the factors relevant to the assessment of the contributions made by this research to our understanding of the changes due to age that take place in colour discrimination.

Firstly. It should be remembered that the samples used subtended a rather large angle 3.5 - 7.5° over the area of the retina, and covered more than the foveal region. Thus, colorimetrically, they belong to the 10° analysis of retinal function.

Secondly, in his hue series he only selected every second member of the original Munsell 100 Hues and thus had only fifty samples to compare with his standards. The colour differences between each of these samples though not as great as they are in the pseudo-isochromatic plates, are still of the order of at least some 3 to 6 N.B.S. units.

Yet, Smith was the first person to show clearly that there was an improvement in discrimination up to 20 and that there was some loss among older subjects. He was also the first to introduce us to the study of the relationship between colour discrimination sex and age, and this marks him as a pioneer in this particular field of colour study; a field which was followed up later by Gilbert and Verriest.

Others who criticised Smith were Alphonse Chapanis and Jeanne Gilbert. Chapanis, in his article entitled 'Relationship between Age, Visual Acuity and Colour Vision' said that he accepted that Smith's results showed a curvelinear function very similar to that found by Galton in his study of visual acuity, but he thought that the data were open to serious question because the colour blind individuals appeared to do as well as, or better than, most colour normal individuals on the tests used. As an explanation for this anomalous result, he quoted the account of one colour blind subject who participated in Smith's study, who told Chapanis that he made his matches entirely on the basis of the texture not the colour of the test sample. Boles-Carenini offered similar criticism.

Perhaps in Smith's defence, it should be said that it is unlikely that the colour blind made their discrimination by texture clues alone, but probably used brightness clues as well. As this was a test of colour discrimination or discrimination as such, the fact that some colour blind people had good scores should not preoccupy us too much as this is now quite a widely known fact. Many anomalous people (i. e. people with deuteranomalous, protanomalous or tritanomalous types of colour vision) can discriminate a 'series of colours' quite accurately. It is only when it comes to colour mixing that their defect is shown.

Gilbert's criticism is perhaps more worthy of serious attention. It confirms the criticism already made in this thesis, that Smith based his conclusions on a limited number of causes and failed to report the exact ages of subjects, <u>so that it is not possible to know exactly when the decline sets in.</u>

Verriest in his paper 'Nouvelles Recherches se Rapportant a'l'Influence du Sexe et de L'Age sur la Discrimination Chromatique Ainsi Qu'a la Signficiation Pratique des Resultants du Test 100 Hue de Farnsworth-Munsell' ('New Researches related to the Influence of Sex, Age on Colour Discrimination') quoted Smith's work as being one in which evidence was obtained from more sensitive tests than pseudo-isochromatic tests and were differential thresholds for hue, saturation and luminosity were tested. He accepted the results which showed some loss in older people, and an increase in threshold for the 20's to 30's. Thus, Smith was the first to give us evidence that colour discrimination improves and is best in young adults while it is less good in both the younger and older subjects.

(b) Criticism of the work of J. Tiffin and H. Kuhn

The main criticisms of this research come from three sources; from the research of Boice, Tinker and Paterson; from that of Kleemeier, and lastly from Chapanis. All subsequent writers on the problem of age and colour vision accepted either partially or wholly these criticisms without adding anything new.

It is necessary here to give some details of these criticisms before any personal assessment is made.

I. Boice, Tinker and Paterson thought the results of Tiffin and Kuhn startling because 26% of the workers between 20 and 25 years of age were quoted as failing to pass the test used by them and this percentage increased steadily with every succeeding 5 year age group until the failures reached 68%. If such a finding could have been verified it would have been of great practical as well as theoretical significance.

They (i. e. Boice, Tinker and Paterson), criticised Tiffin for failing to name or describe his 4-plate colour vision test. They corresponded with Tiffin and found that this test consisted of 4 colour plates which were included in the 'Keystone Telebinocular' visual testing apparatus. The validity of this abbreviated 4-item test was not given, but they thought it was probably low. Their own conclusions about this were based on the results of Foster's 'Comparative Study of Three Tests of Colour Vision' using the Ishihara, the American Optical and the Yen-Sen Tests. It was found that the Ishihara and the Americal Optical Co. Test yielded results similar to each other and they claimed that a 26% failure in the 20 to 25 age group was suspect because it was not in line with the usual percentage results.

Tiffin was also criticised for failing to describe adequately the sample tested (e.g. intelligence, education, occupation and socio-economic status were not held constant). Without these controls it was entirely possible that the results reported might have been due to any one, or any combination of those factors, rather than to increased age per se. The number of workers (7,000) tested was considered impressive, but it was thought that more rigorous controls of all possible variables was necessary.

However, in a footnote, it was admitted that Tiffin and Kuhn were cautious in their claims and had themselves said that failure in the test did not necessarily imply colour blindness. Boice, Tinker and Paterson complained that a similar caution was not included by Tiffin in his Textbook which is much more widely read than his article. Thus many readers accustomed to thinking in terms of colour blindness rather than of normal colour vision would undoubtedly interpret the results in the report as indicating that there is a striking increase in colour blindness with advancing age.

II. Chapanis Criticism covered essentially the same points made by the previous writers, that is, he questioned the validity of the test used by Tiffin and made the same mistake in thinking that the 26% incidence of failures in the 20 - 25 age group referred to the incidence of 'colour blind persons'. ......"this high proportion of failures......is completely at variance with the results normally obtained from other colour vision tests".

He added that some of the other data obtained by Tiffin and Kuhn, in particular the positive correlation between occupation and colour vision score, suggested that this test might be measuring factors other than colour vision (for example, intelligence, education or, more generally, test sophistication).

Lastly, Chapanis pointed out that, in his discussion, Tiffin himself admitted that he was not satisfied with this particular test for segregation of industrial employees, as nothing was known about this device as a colour vision test.

III. Kleemeier's main contention was that unless the acuity variable was excluded from any study of colour vision and age, the results obtained might be invalid, as for example in the case of Tiffin and Kuhn.

The author seemed to think that "acuity" as such is the expression of some specific visual function possessed by each subject. In this he is mistaken for in this way the concept of acuity refers to some 'abstract' ability. Acuity is dependent on many factors, some external such as the level of illumination, and the size of target, and some 'internal'. Of these the state of rod and cone function is an important factor. Colour discrimination and acuity (i.e. discrimination of 'black and white') have a common receptor system, and in this sense, colour discrimination forms a part of the complex function that we term acuity. What Kleemeier did not realise was that losses in colour discrimination are the first signs of the changes that must take place in the visual system. Such changes indicate the very small changes that are taking place when that system is ageing. Thus, colour discrimination and acuity are not exclusive terms.

It was noted previously that the work of Boice and that of Kleemeier were undertaken with the express purpose of ascertaining the validity of Tiffin's results and to examine the aspects of the conclusions. A large number of misreadings, comparable in magnitude to those we obtain from genetically determined red-green colour defects, was expected, as a proof of deterioration of colour discrimination with age. Yet the task of perceiving the figure and the background in a pseudo-isochromatic test, as opposed to a test measuring the extent of colour discrimination as such, required quite different abilities. The plates are so constructed that overall differences in figure and background are of the order of 20 to 35 N.B.S. units (i.e. about 150 j.p.s.) which is much greater than what was expected from using the Keystone Test which Tiffin and Kuhn employed. Misreadings such as M. Boice found in her study of the older academicians can only be the result of gross changes in either the lens structure or the retinal or neuronal structure of the visual system. Losses involving 15 or 20 misreadings on say the Ishihara plates, indicates a loss in colour vision which nowadays would be termed an acquired dyschromatopia arising from central scotomas. Of course, around 1950 little was known about the work done on such colour losses by the 19th century German scientists, and it was some 5 years later before new interest in this field sprang up.

In another respect, the conclusions reached by Boice and Kleemeier were rather odd. Though the actual scoring cannot be found from the papers published on these pseudo-isochromatic tests there must have been some evidence for changes in performances even if on a smaller scale. In the light of the results obtained in this research when using the Dvorine plates, it is surprising that even Chapanis made no attempt to analyse the data present, since he used a battery of pseudo-isochromatic tests which included the Dvorine. Such evidence must have been present, yet was either neglected or ignored.

From the historical point of view, Tiffin's and Kuhn's case is a very interesting instance in the development of scientific thought. It illustrates aptly how difficult it is for a new explanation, or new account of phenomena to take root, while orthodox explanations prevail. In the forties, it was still believed that only two types of colour discrimination existed - colour defectives and colour normals and in this last group there was no room for differentiation. It was assumed that all non-defectives possessed the same type of ability. On this basis it was very difficult to accept the conclusion that already at 25 years there could be some with less than perfect colour vision. It was only in the fifties that the field of individual differences in colour discrimination was explored, and new concepts such as 'Colour Weak', 'Deviants' etc.

introduced by Pickford, slowly became acceptable.

The idea that minor variations in colour discrimination could point to variations in general physical and health conditions was too novel to be accepted. True enough, the German scientists of the nineteenth century described cases where colour discrimination losses were greater in subjects who were not genetically determined colour anomalous, but such cases were thought of as occurring in only a small minority of people - such as the oddities found in ophthalmological clinics. Though the idea that acquired dyschromatopsias could exist prior to senescence was first studied about the middle of this century by people such as Zanen (1953), Francois and Verriest (1957), it is only now that it is accepted that all ocular diseases can give rise to absorption losses, and that lesions in the optic system can lead to losses in yellow-blue and red-green discrimination.

Historically speaking, therefore, for Tiffin and Kuhn to say in 1942 that perhaps Colour Vision was a barometer of general health was absurd, for such a relationship was not considered to exist at that time.

Lastly, this was still a time when any test of colour discrimination was looked upon as a test measuring only the variations of the degree of colour deficiency in major colour defectives. Thus, finding deficiencies as measured by the 4-item test used by those two authors was synonymous with finding people whose performance was similar to that of major colour defectives. This was still a time when proper measurement of the stimulus had not been thought necessary, for, in practice, only tests that could classify people into two groups were considered to be proper tests of 'colour vision'. It is a pity that Tiffin and Kuhn did not give any account of the physical characteristics of the test they used, as it would have been simple (now at least) to ascertain how near they were, even if accidentally, to using a very sensitive test.

Later in this thesis it will be shown that even colour naming tests can indicate 'losses' in colour discrimination, provided it is known spectrophotometrically what colours are used, and where in the 'colour space' the sampling is made.

(c) Evaluation of other contributions, i. e. Janouskova's Warburton's, Brown's etc.

It is surprising that in Brown's work with the anomaloscope (metameric matching), no significant differences were found in mid-matching points between adults and children.

Boles-Carenini's contribution is consistent with the conclusions offered in this thesis, but one wonders why a more elaborate analysis was not attempted. For example, he made no distinction between the idea of matching ranges and mid-matching points and used a method of recording and scoring results which would not be acceptable to most users of anomaloscopes. Each subject himself selected ten matchings which were then analysed as though they had come from ten different subjects. This method did not permit the determination of the full range of matches that might exist i. e. adjustments made by the experimenter might have produced many matches which did not occur spontaneously when the subject was merely asked to find a match. This is especially critical for the older subjects, because variations in matching ranges are a sensitive measure of ageing.

In the case of the work of Janouskova, the use of large numbers is commendable, and the ingenuity shown in testing aphakic subjects is of interest. It is a pity that testing was done only on pseudo-isochromatic plates, and rather crude tests of colour discrimination, and it is difficult to understand why the author did not make a similar analysis of the performance for those in say the 15 to 55 age groups.

Warburton's study was most ingenious, and the results obtained within the limits of the experimental set-up. What is rather puzzling is that this research worker found it impossible to assess the 'colour temperature' or the 'number of sample' that should be chosen by the 'standard observer'. All the necessary information appeared to be available.

2 : 8 SUMMARY AND CONCLUSIONS

The main points of each of the researches so far described are enumerated below, with emphasis on material relevant to the evaluations of this present research. This is followed by a chart, summarising the general position as it was in 1951 - 58.

(a) H. C. Smith claimed that ability to discriminate colours by matching improved from 6 years of age, reaching a peak at 25. There was no noticeable decrease in this ability for those in the 30 to 64 age range and only the very old showed any decline in ability. Sex differences were found only in the 5 - 11 age group, where girls were better at matching hues and saturations. The most difficult colours to match were found to be blues, and yellows were easiest. No correlation was found between colour matching ability and mistakes made on the Ishihara.

Differences in colour matching were considered to be due mainly to learning factors.

(b) Tiffin and Kuhn stated that besides deficiencies of the red-green type, 4.4% of the testees also showed yellow-blue losses. There was a significant correlation between high job performance and good performance on colour vision tests. They found that colour vision showed a marked reduction with increased age, noticeable even in the 26 - 30 age group. The classification of people into colour blind and normal they considered insufficient, and thought that young people with very good colour discrimination should be chosen as workers, since this ability decreased with age.

(c) Boice, Tinker and Paterson found that there was a greater incidence of 'colour blind' people than the usual 7.5% quoted for the general population, but only among the over 60's. They recommended that large scale research should be initiated where intelligence, education and occupation levels were controlled.

(d) Chapanis in both papers (1948, 1950) claimed that there was no significant correlation between age and colour discrimination. Only examinees who were 14 years of age or under gave poorer than average scores. However, he added 'these zero correlations do not show that such relationships do not exist, but rather that pseudo-isochromatic tests are perhaps too crude to elicit them'.

(e) Kleemeier thought that poor colour discrimination in the old was caused by deficiencies in visual acuity and not by 'colour vision' as such.

(f) Janouskova's conclusions were that children are born with the basic physiological colour mechanism, but that maturation processes were at work, as evidenced by the progressive development of colour discrimination until 12 years of age. It was possible that in girls, colour perception developed more quickly than in boys. Among older people, from 51 years onwards, there was a greater incidence of defective colour vision, reaching as much as 26% for those in the 81 - 85 age group.

Colour naming and colour recognition was found to be better in subjects with post-operative aphakia than in those with incipient cataract. A 'clouded' lens noticeably altered colour perception and caused deterioration in certain parts of the spectrum, but not all the deterioration seen in senile groups could be attributed to lens changes alone. It was suspected that it could also be caused by changes in the receptor system of the retina or in the structure of the cortex.

(g) Brown's conclusions were negative, no evidence of any significant difference in the colour vision of school children and that of adults was found either when using the anomaloscope or the Ishihara plates.

(h) The variations due to age in normal colour vision found by Warburton expressed as variations of colour temperature of an illuminant 'B' (i. e. 5000°K) showed differences between the young observers of 16 - 25 years of age who matched it near illuminant 'C' and subjects in the 56 - 65 age group who matched near illuminant 'A' (i. e. the variations were between 3500°L to 7500°K). Warburton claimed that his evidence gave concrete proof of increased macular pigmentation with age. Because of the large variations among his observers, he considered that standardisation of light sources for colour matching alone would not be sufficient to eliminate differences in practical matchings.

(i) Boles-Carenini stated that using the Nagel anomaloscope, differences among his observers came mainly from the fact that the 20 - 29 age group had different colour matching ability from the rest of the age sample. Only for the 60 to 79 age groups were the means of the matches significantly worse than elsewhere. This significance was mainly attributed to the fact that the mean of the Rayleigh equation for those old people was displaced towards the green part of the equation. His final conclusions were that there was no relationship between age and chromatic sense in subjects from 20 to 60 years of age. However, after that age the possibility of dyschromatopsias occurring would increase even in subjects possessing normal colour vision. The possible pathogenic mechanism of such anomalies might be attributable to senile sclerosis.

(j) Quellette considered that young people had better discrimination than older people. There was no sex difference, and no one colour showed a more rapid decline in discrimination than any other, though green appeared to be the most difficult colour to match accurately.

(k) Gilbert on the other hand, claimed that the ability to discriminate blues declined more quickly than the ability to discriminate greens, whereas the decline in this ability for yellow and red seemed less affected by age. Females were found to be superior to males in total scores, though this varied for the individual colours, and for the different age groups. In this colour test children did not perform as well as adults, although the reason for this was not clear. It was considered that it might be due to developmental factors. Those in the 20 - 30 age group scored significantly better than those in all other decades.

Generally speaking, the ability to match colours developed in childhood, reached a peak in the 20 to 30 age group, and from then on, showed a steady decline to old age. This decline was gradual, so that only differences between alternate decades were statistically significant. Losses found were attributed to the filtering effect of the progressive yellowing of the lens with advancing age. It was thought, when selecting colour workers, that it would be desirable to ascertain that, initially, they had very good colour discrimination.

Author	Year	Subjects No.	Subjects Sex	Age Range	Test Name	Test Type	Colour dimensions	subtens	Conclusions
Chapanis	1950	574	M & F	5-79	Authors, Boström & Kugelberg, Ishihara (9th ed.) Dvorine (1947) A. O. 1940	Dichotomous qual.	R - G	40' - 8^{o}	Age has no effect on colour discrimination
Brown	1950	770	B & G	5-15	Ishihara (10th)	Qual. Dichotomous	R - G	40' - 8^{o}	
		120	B & G	5-15	Pickford	Anomaloscope	R-G & Y-B	40'	
Kleemeier	1952	128	M	67-85	Ortho-rator	Dichotomous	R - G	40' - 8^{o}	
Boice et. al.	1948	236	M	20-60+	Ishihara (7th) A. O. (1940)	Qual. Dichotomous	R - G	40' - 8^{o}	Deterioration only
Boles-Carenini	1954	170	not given	20-79	Nagel	Anomaloscope	R - G	prob. 30	after 60 years of
Quellette	1955	56	M & F	20-30 & 75-85	C. A. T.	Aptitude Test	R, G, Y, B.	3^{o} - 8^{o}	age
Smith	1939	199	M & F	5-65+	Munxell pop.	Colour Matching	Y, G, B, P. R.	3^{o} - 7^{o}	Growth of discrimination till about 25 years
Janouskova	1955	3450	B & G	3-12	A. O. (1942) Ishihara	Dichotomous	R - G	40' - 8^{o}	of age. Deterioration only after 60 years of
					Silks	Naming	R, Y, B, G, V.	?	age.
		3000	M & F	51-85	A. O. (1942)	Dichotomous	R - G	40' - 8^{o}	
Tiffin & Kuhn	1943	7000	M	20-70	4-item test	Discrimination ?	R-G & Y-B	prob. 5^{o} - 8^{o}	General deterioration or change from 25
Warburton	1953	247	not given	16-65	Bradford	Colour Matching	White	4^{o} - 5^{o}	years of age upwards
Gilbert	1957	355	M & F	10-93	C. A. T.	Aptitude Test	R, G, Y, B.	3^{o} - 4^{o}	Growth of discrimination till 25 years then general decline

3. ANALYSIS OF RESULTS INCLUDING COLORIMETRIC AND PHOTOMETRIC SURVEY

The main conclusion to be drawn from the historical review is that studies prior to 1958 do not agree about the age at which the onset of colour deterioration is found, about the extent of such deterioration, and some even deny it existence altogether.

These discrepancies arose from three sources :-

Firstly, from the difficulty of defining what the term 'colour vision' really means.

Secondly, from the assumption that tests for detecting congenital colour vision defects, and tests for colour discrimination, as well as tests of the anomaloscope type, all deal with the same attributes of colour vision.

Thirdly, from inadequate experimental control, either due to there being too few subjects in certain age groups, or from insufficient statistical analysis, or finally from sampling error.

In an attempt to clarify some of the issues raised by the first and second sources of discrepancies, it is necessary at this point to discuss what is meant by the terms 'colour vision' and 'colour vision tests'.

3 : 1 COLOUR VISION

(a) Definition

The concept of colour vision as distinct from the simple linguistic meaning is rather a complex one, and implies a great variety of experiences on the subjective side. This concept has grown along with our understanding of visual processes. At first, as in sound, any analysis made was of subjective accounts of the awareness of the coloured world around. When more objective tools of analysis were available they were still expressed in terms of personal experiences and had very little reference to the nature of the stimulus. Only very recently has an objective description in terms of the physical attributes of the nature of the stimulus been possible, and this in turn, has transformed the terms of reference in our description of the phenomenon of colour. It is no longer man orientated alone, but must take account of the precise nature of the stimulus and must accept a 'Standard Observer' as a criterion of perception.

Thus is colour vision, we are constantly using a multitude of concepts that belong to the fields of physics, psychology and psycho-physics. Another factor must also be included. The actual process of colour vision depends not only on the nature of the stimulus and the awareness of it, or its perceptual aberration by the observer, but depends also on the innate or acquired neuro-physiological make-up of the subject. Therefore, in the study of colour, three fields of inquiry must necessarily be constantly considered, since in the last analysis, colour vision as such is the by-product of all these elements. These are - the field of colour perception proper; the neuro-physiological make-up of the perceiver; and lastly, the nature of the stimulus.

What is meant by the first of these three elements (i.e. colour perception)? It is only a part of visual perception, and therefore any definition of colour perception is also a definition of the former. Thus, if visual perception is defined as the intergrated conscious response to a total visual stimulus situation, it must also be added that it is a response modified or interpreted in terms of stored remnants of past experience which are brought into use in that particular situation. Various aspects of visual perception, such as colour, form, movement, size, depth may be abstracted. In this sense, 'colour perception' may be specified in terms of the dimensions of hue, saturation and brightness. Colour perception is therefore equivalent to the term 'colour' which implies awareness of the non-spatial, non-temporal attributes of visual perception. Perhaps it should be added that in using the term 'colour sensation' we refer to the least complicated case or prototype of colour perception in the sense that it involves less interpretation in terms of past experience. Yet when it occurs, colour perception always arises in various spatial and temporal contexts, which comprise the modes of appearance of colour. As far as colour vision tests are concerned, three types of modes should be mentioned - surface, volume and film modes. Colour perceived as belonging to a surface is said to be perceived as surface mode. The majority of colour vision tests, and all pseudo-isochromatic plates such as the Ishihara and Dvorine and the Farnsworth-Munsell 100 Hue Tests, belong to this category. Colour perceived through the bulk of a uniformly transparent substance is aid to be perceived in the volume mode. When film mode is referred to, the implication is that the stimulus object is not distinguishable as an object, as for example in the situation where colour seems to fill the space behind an aperture. This happens when colour dots are viewed on the Pickford Anomaloscope. It should be added here that film mode of appearance is considered to be the simplest of all the modes in the sense that it is the least removed from pure sensation, whereas, in the case of 'surface mode' the complex attributes of full visual perception, such as form, size and especially the context in which the colour surface is perceived (such as glossiness, texture, glitter, etc.) can never be estimated.

So far, the broader aspects of the term 'colour vision' have been outlined briefly. It is an act of perceiving a change in the visual field.

Now the more specific meaning of colour vision will be discussed, particularly in relation to the problem of age and colour vision. This can, at first, be defined simply as the 'awareness of differences in colour stimuli'. Such an awareness can nowadays be quantified, and because of this, the problem can be studied by scientific methods instead of being a topic for purely philosophical curiosity, as it formerly was.

Whenever a phenomenon can be quantified, by implication, it must also be possible to measure the extent of any differences. Thus, measurement of differences in colour stimuli becomes a very important activity, even a prerequisite to our understanding of the entire nature of colour vision.

The measurement of the differences between colour stimuli is only meaningful, if we know the limits of colour change awareness for a given human observer. It therefore becomes necessary to study and define the extent of colour discrimination capacity human beings possess, for 'colour', as it is to be understood in the context of this thesis, is only meaningful if by that we refer to the end product of a complex process of perception which finally becomes the <u>conscious awareness</u> of a difference in the environment. In this sense, when we talk about red, green, yellow colours etc., we refer to a mental act.

The extent of conscious awareness is, of course, dependent on the state and type of receptor mechanism an individual possesses, and on the physical nature of the stimulus.

Thus, differences in colour stimuli are only meanin gful as long as they can be perceived, or are capable of being perceived, by an observing subject. The study of the extent of the differences present in the stimulus is limited by the extent of the individual's colour discrimination capacity and, therefore, all discussion must be centred upon the measurement of the extent of colour discrimination that we are capable of experiencing.

In most studies on the problem of age and colour, the fact that colour discrimination is a capacity we all have, was taken for granted. What was aimed at was finding what change in, or loss of, this ability was due to the age variable.

The measurement of the extent of colour discrimination has been achieved by many devious and straightforward ways, from the assessment of how various people describe differences of colour experience, to the refined measurement of awareness of colour difference as studied by means of colour mixing instruments, like the Helmholtz Spectralmischapparat. In the first example, complex stimuli had to be recognised, while in the other, highly abstract aspects of colour phenomena such as discrimination of hue, lightness, or saturation, were studied.

In the pages that follow, the description will be confined to those aspects that are important for measuring the extent of colour discrimination, or its variations among individuals, and a discussion will be limited to problems relevant to the methods of study used by research workers who studied changes in colour vision with age.

(b) Measurement of Colour Discrimination

The measurement of colour discrimination provides not only indices of the limits of normal human vision, but also yields us considerable information about the differences existing from person to person in this dimension.

Such measurements may be specific in the sense that there may be a difference in perception of either hue, saturation or brightness, or they may be general in the sense that what is perceived is an overall difference in colour. Another approach to this question is that of measuring the degree of 'fineness' of colour discrimination on the one hand, and the extent and type of colour vision defect on the other.

To ascertain the 'fineness' of colour discrimination, various psycho-physical methods are employed - those measuring absolute and differential thresholds. These thresholds can be used for specific studies of the three basic colour response dimensions already mentioned (i. e. hue, saturation and brightness). Thus, work has been done upon the absolute and differential brightness thresholds, upon panchromatic and differential thresholds, and upon the absolute and differential thresholds of saturations. These studies have gone hand-in-hand with researches into the pecularities and characteristics of colour vision defects. The first type of research, of course, has had a long tradition and a more methodological approach. It was usually done by people who had a good grounding in physical optics, since, of necessity, such studies required complicated instruments with optical devices, if the researcher were to be able to obtain the degree of purity or precision required to isolate any of the given attributes of colour being studied.

In studies on colour vision defects, less stringent conditions were needed, as what was required was the measurement of a gross variation in colour perception. Less exact tests, instruments or procedure could yield all the diagnostically necessary results. Thus a more varied group of workers, including ophthalmologists, chemists, general medical practitioners and even psychologists, has been engaged in such studies. This, in turn, led to a profusion of colour vision tests devised on less stringent 'a priori' conditions. Most of the pseudo-isochromatic tests were designed on empirical grounds employing either colour blind artists to devise a test (Stilling test) or by a trial and error method employed for finding a combination of colour dots which the colour defectives would confuse (e.g. Dvorine test).

It is this lack of precise knowledge of how the tests were constructed (i.e. there is no account of the photometric or colorimetric nature of such tests, or of attributes of colour they are testing), that led to the rather contradictory results obtained from studying colour vision along the age continuum. Of course, other factors such as lack of proper procedure and application of tests, too few subjects in age groups, lack of true randomness, etc., have all added to this.

At this point, our attention must be focussed on the fundamental aspects of the nature and construction of colour vision tests. Most of the tests used were rather less than sophisticated in terms of their physical construction or optical make-up, since in studies involving a large number of subjects, easily administered tests must be employed in order to avoid other factors of the test situation bedevilling the analysis of results necessary to assess colour discrimination as such.

I. Tests of colour vision - In the researches on colour vision and age already discussed, a variety of colour vision tests were used, such as pseudo-isochromatic plates, colour aptitude tests, colour discrimination tests, anomaloscopes, etc. Analysis of such tests can be approached in many ways. For example, there are -

a) tests of inborn differences in colour capacity;

b) tests in which the complexity of colour vision involvement is tested;

c) tests where complexity of colour stimuli is used;

(i) with narrow or broad transmissions or reflectances,

(ii) involving either metamerism or isomerism,

d) considerations of the subtense that will imagine on the retina.

Perhaps it should be added that in the analysis to follow, the discussion will be extended to include tests of colour vision used in research on age and colour vision since 1958.

(i) The tests most frequently used in studies of colour vision and age were tests which attempted to distinguish the inborn differences in colour perception. This, of course, is not always clearly achieved and, therefore, it is necessary to indicate whether a test of inborn differences does not in fact measure acquired colour discrimination capacity also, since in the last analysis, ability to discern must depend on two aspects :-

(i) the native or inborn, which will include intelligence, the state or type of physiology of the colour vision system exhibited by the individual, and other inborn perceptual factors, and

(ii) acquired factors such as learning, and learning of perceptual skills developed during training.

There are many tests of colour perception which measure inborn colour capacity. These tests are so designed that training supposedly has little effect on the performance on them. Broadly speaking, there are three categories of such tests - the first and second are called respectively, dichotomous and qualitatively diagnostic tests, and the third type is called a quantitatively diagnostic test. Of the well-known dichotomous tests, the Holmgren-Wool test is the best example. Here, a subject has to sort chromatic wool samples into hue groups. Other dichotomous tests are the American Optical Co. Pseudo-isochromatic plates or the A.O. Test used by Janouskova. There is also the Farnsworth dichotomous test for colour blindess or the so-called Panel D.15 which is a test where colour chips are arranged in an order which, for normal colour vision, is according to the hue, but for dichromatic vision, is according to the saturation. The tests so far mentioned are the simplest colour vision tests.

The second type of colour vision test is the qualitatively diagnostic test so named because it is designed to classify the type but not the degree of defective colour vision. To this category belong the world-famous Ishihara and Dvorine Tests. In a sense, these tests are dichotomous since they tell us whether the person is normal or defective in respect of colour vision, but they also attempt to give us a qualitative diagnosis though the extent of defect is not specified, - only the type. For example, in these two tests there are plates (in the Ishihara 4 and in the Dvorine 2) separating protanopes from deuteranopes, but these plates do not tell us the degree or extent of red-green defect. At this juncture we ought to point out that most of the tests so far mentioned are tests of red-green vision only.

The Sloan Color Threshold Test is a quantitatively diagnostic test that uses the recognition of chromatic lights to measure degrees of red-green deficiency without distinguishing the type of defect. It is essentially a threshold colour naming test. Another well-known test is the American Optical Co. Hardy-Rand-Rittler Pseudo-isochromatic Plate Test. This is qualitatively diagnostic and has been standardised to distinguish both the type and degree of red-green and yellow-blue defect. Lastly, there are the anomaloscopes which give a qualitative and quantitative analysis of colour vision defects. Most of the anomaloscopes test red-green vision only, utilising the so-called Rayleigh equation. The most famous and most widely used of these anomaloscopes is the Nagel.

The anomaloscope used in this study and for the purposes of this research is the Pickford Anomaloscope. This instrument tests not only red-green discrimination, but also yellow-blue and violet-blue-green.

(ii) The tests so far classified are tests of native colour vision. The next group of tests includes the so-called tests of 'colour discrimination' which tend to overlap tests for 'acquired' colour ability.

Perhaps it should be added that there is no test that is completely unaffected by ability acquired through training, but some tests of colour discrimination tend to be more affected by training than others. The 100-Hue Test for example, would be classified here as a test of native ability tending to give results least affected by training. It is also called a perceptual test, as opposed to a memory test alone.

The Farnsworth-Munsell 100-Hue test is used for several purposes :-

1. It identifies varying degrees of perceptual hue discrimination among people with normal or defective colour vision.
2. Though it is not specifically designed as a dichotomous colour vision test, it may be used with others to corroborate evidence for colour deficiency. It can discriminate tritans as well as protans and deutans. The test makes it possible to classify people with normal colour vision into three groups, superior, normal or low in respect of hue discrimination.

(iii) Lastly, there are the tests designed to measure acquired colour skills, developed through training in making various kinds of colour judgments. Tests of acquired colour skills draw on inborn factors, but the scores are influenced by the amount of specific training the person has received. The test of acquired colour skill referred to here has been developed for the Inter-Society Color Council and is called the 'Color Aptitude Test'. It requires the subject to discriminate small differences in saturation and this depends on experience as well as on the inborn capacity to make such discrimination.

II. Complexity of Visual Involvement - This depends on the amount of perceptual involvement that a given test requires. Thus, for example, the perceptual involvement in most of the pseudo-isochromatic tests is extreme and rather complex. Not only are the subjects required to know numerals (usually Arabic) or recognise shapes (usually geometric), but also the state of adaptation of the subject is important. The peculiar structure of the material on which the test has been printed is a very important factor leading to additional clues being available outside the colour dimension. In colour discrimination tests such as the 100-Hue, the concept of 'colour series' (i.e. that colours gradually change from red to orange through yellow to yellow-green) is an important one which depends on maturation as well as on the physiological nature of colour vision. On the other hand, there are tests that involve little learning, require an average amount of intelligence, and, in the main, depend on the state of neuro-physiology of the visual system of the subjects tested. Anomaloscope tests fall into this category. In this instrument the colour match achieved is relatively independent of the state of adaptation. Though the colour naming will change, the match on such tests remains the same through a large gamut of illumination levels, provided that the general luminosity value of the matching field is above the scotopic level. Thus, such an instrument can even be used at mesotopic as well as photopic levels of illumination without essentially altering the relative scores. This is not true of tests of the pseudo-isochromatic type, where adaptation to the level of the illumination and colour temperature of the source are important factors.

III. Complexity of Stimulus -

(i) Colour vision tests can also be classified in terms of the complexity of colour stimulus involved. For example, in pseudo-isochromatic tests, and in tests of colour discrimination, (both of which use surfaces) in physical terms the stimulus is rather a complex one. The resultant colour not only depends on the peculiarity of the surface in terms of absorption, be it a broad or narrow band, but also depends to a large extent on the type of illumination. At any given point in a test, that is at any plate, the figure and background are usually composed of 2 to 6 different coloured dots and these have different reflectance characteristics.

The anomaloscope on the other hand, is a test that is independent of the type of illumination because the matching surfaces are illuminants in themselves, and in this sense we actually match two spectral lights. Anomaloscopes vary in the complexity of actual utilised energy. In the Nagel for example, there are simple monochromatic lights, whereas in anomaloscopes using filters, rather broader band transmissions are employed. In the Pickford anomaloscope, spectral filters are used. These are not as narrow in transmission as the stimuli used in the Nagel, and, in fact, have transmission bands of 10 to 15 mu for each primary used.

(ii) Colour vision tests also differ in the degree of metamerism or isomerism involved in the construction of the test.

The term metamerism was first adopted by Wilhelm Ostwald to describe the phenomenon exhibited by two or more surface colours which appear to be alike under one colour of illumination such as daylight but did not match under another type of illumination such as an incandescent lamp. Although the term was coined only about 40 years ago the phenomenon was by no means unknown. Studies of extremely metameric pairs in which mixtures of two parts of the spectrum are set up to match in colour another two parts spectrum mixtures have yielded us the most valuable knowledge about the properties of the average normal eye. In this sense, then, metameric pairs as they are called, have to be seen as stimuli having the same colour (i.e. subjectively) but they are different in spectral composition. The essential idea here is that we are dealing with two levels of explanation, one involving the physical properties of the stimulus and the other the sensation that is produced in a human observer. When colours look alike whether they come from a surface or from a source there are two basic reasons :-

(1) Either they are physically or spectrophotometrically identical i.e. the energy distributions in sample "a" and sample "b" are exactly the same). This is called an 'isomeric match' that is, one in which both surfaces or lights look alike because they have the same physical characteristics.

(2) There are, however, surfaces and lights which are dissimilar in physical characterisation and these sometimes do match. These matches arise where two physically different stimuli (as far as spectral composition is concerned) result in an identical sensation. This we call metamerism. The property exhibited by

metameric pairs is such that if the members of the pairs are of very different spectral compositions they are said to exhibit a high degree of metamerism. If, on the other hand, there are only slight differences then the metamerism is of a lesser degree.

In an anomaloscope the colours to be matched are essentially in a metameric relationship, that is, though the two fields have the same tristimulus values, they are spectrally different. Thus, what has to be ascertained is the degree of metamerism that can be accepted by an individual as a match. This form of colour vision testing is one of the most sensitive and depends entirely on the neuro-physiology peculiar to the visual system of a particular individual.

In other colour vision tests where figure and background had to be distinguised, the differences are of the isomeric type. That is, two or three such colour stimuli are either isomerically different or identical. If they are identical, the spectral composition of the surfaces is exactly similar in every way, and consequently these tests are cruder in terms of stimulus characteristics. If differences are introduced, they are generally either in lightness or saturation, but they might also be metameric (e.g. in Warburton's study).

IV. <u>Subtense - 2^{o} or 10^{o}</u> - Colour vision tests also vary in terms of the subtense they image on the retina.

Some confine themselves to a 2^{o} angle while others are of 10^{o} or more. In most anomaloscopes, for example, this can be varied. In the Nagel the subtense most often used is one of 3^{o}, while in the Pickford anomaloscope the subtense is rather small at 40 mins. In tests using surfaces, we come across a great variety of subtense. For example, the individual dots that make up pseudo-isochromatic plates when viewed at 30 to 40 cms. may subtend from as little as 40 mins. for the Tokyo Medical School Colour Vision Tests, to the $1\frac{1}{2}^{\circ}$ subtense found in the dots on the Farnsworth Tritan type of test. Again it is very difficult to know what the eye actually discerns in such tests, whether each dot is perceived individually, or whether the perceptual act is in terms of the whole figure against the whole background. If the latter is the case, the subtense would then be of the order of 5 to 10°. In the 100-Hue Tests for example, each individual cap subtends $2^{\circ}\ 10'$, and in the Colour Aptitude Test the size of the individual disc is approximately 35 mm by 45 mm leading to a subtense of 3 to $4\frac{1}{2}^{\circ}$.

These are very important considerations because colour vision up to 2° is essentially different from that of over 2° subtense, and therefore when assessing the results of a test, it is essential to know exactly which part or how large a part of the retina has been stimulated. There are, of course, individual differences for a 2°, and for a 10° subtense, but the changes shown in age, using a 2° subtense are not exactly the same as those shown when using a 10° subtense. The deterioration noted when using a subtense of up to 2° follows a different pattern and depends on different aetiological causes than the deterioration in discrimination in the parafoveal region of the retina. While losses in the parafovea and periphery of the retina are due mainly to pathological causes, losses of sensitivity in the fovea alone arise from two factors, first, aging (i.e. senility) and secondly from pathology of the visual system.

3 : 2 TESTING - DESCRIPTION OF POPULATION AND PROCEDURE

Over 550 people of both sexes were tested in different parts of Scotland, and in view of the findings of Vernon and Straker (1943) on the distribution of colour blind men in Great Britain, a more detailed account of the geographical distribution of the population tested is given.

The main localities where testing was done in the West were Glasgow, Dumfries, Dunoon, and in the Centre and East, Wishaw, Motherwell, Coatbridge and Edinburgh.

(a) Geographical distribution of population

I. West of Scotland -

Glasgow - Here testing took place in four different places : the University, St. Joseph's Home for Old People, Foresthall Geriatric Unit, and a Church Hall in East Glasgow.

The University. In the Department of Psychology about 23 students of various ages were tested. Some of the University's cleaners and servitors were tested and also several friends and acquaintances. However, most (i.e. about 40) of the people tested here were passers-by who were asked to participate in colour vision tests. Pupils of Motherwell Junior Secondary School (altogether about 58 boys in the age group of 12 to 15) and another 14 people from the Motherwell area were also tested in the department.

St. Joseph's and Foresthall Homes for the Old. These are homes for old people in Glasgow and the subjects here were chosen on the advice of the medical staff and the Mother Superior. This group consisted of people in the 50 to 90 age group. All were chosen because they were subjects who could be expected to co-operate and were more active than other inmates. Of the two groups, those who came from St. Joseph's Home were better off economically and had the better social background, but as a whole these people were not institutionalised in the accepted sense. They were active, and able to leave the home to do their shopping, visit their relatives, or live for the weekend at home. About one third of these subjects were tested twice. Testing time of necessity was prolonged taking a minimum of one hour, and this had to include many breaks for relaxation. Altogether 76 subjects were tested in these places.

Parish Hall, Abercrombie Street. Here about 70 subjects were tested, and of these, 33 were adults between 25 and 80 years of age. Most were random passers-by who were asked to participate in tests of colour vision. In addition to this group of adults around 40 local children of 10 and under were tested.

Dumfries - At the Crichton Royal, a mental hospital in Dumfries, a group of 60 patients, consisting of 30 schizophrenics and 30 manic-depressives, were tested. All patients were under the age of 60, and of these, only 20 were chronic (not aged) patients. An analysis of the Means and Standard Deviations of the red-green mid-matching points for the 43 patients included in the final calculations, compared with the scores obtained on the 16 to 45 'normal' age population, shows no statistically significant differences at the 5% level. Such a result permits the inclusion of this special group of subjects in our age study.

Fifteen others at this Hospital among them doctors, almoners, nurses, gardeners and visitors were also tested.

Dunoon - Here a group of 17 subjects of different ages starting from the 30 plus age group, were tested.

II. Central and East Scotland -

Coatbridge, St. Mary's Junior Secondary School - Tests were given to 37 girls in the 12 - 15 age group, to about 5 of the primary children of this school, and in addition, 17 others were tested, among them staff, cleaners and people from the neighbouring area of Whifflet.

Edinburgh - Altogether 116 subjects were tested in Edinburgh, all age groups being represented in this study, but the majority of them came from the 26 to 50 age groups. Many of the subjects were members of the works department of the University, while others were members of a club and about 11 were children under 10 years of age.

(b) Testing Procedure

In every place where testing was done, a room that could be blacked out was used so that all patients tested on the anomaloscope were tested under identical conditions. Testing by means of the Ishihara plates and the other pseudo-isochromatic plates was done under a standardised lamp (colour temperature 6,000°K), and was carried out either in the same dark room, or in another place where daylight was part of the general illumination in the room. However, the testing bay was always arranged in such a way that the main part of the illumination came from the light source provided.

The sequence of presenting tests was as follows : Firstly, the subject was tested on the Ishihara Test, then on the Dvorine, and after this was given the Tritan Plates. In some cases subjects were then asked to do the nomenclature test, and lastly testing on the anomaloscope proceeded. Here the red-green equation was tested first, then the yellow-blue and finally the violet-blue-green. Prior to testing on the anomaloscope a 15 minute period was allowed for pre-adaptation. Further details of procedure and scoring method will be given when the tests and their results are discussed.

It should be mentioned here that, although altogether 550 subjects were tested, only 400 are used in the analysis of results obtained from the anomaloscope. Care was taken to omit all those whose results for some reason or other were suspect - for example, all students were eliminated from this analysis because of a general criticism that the majority of studies of sensory discrimination are made on student populations. Secondly, of the 76 people tested from the Homes for the old, only 41 were included in the anomaloscope study, where it seemed reasonable to assume that the results represented the actual sensory discrimination and not other factors, and in the same way, only 43 of the 60 patients tested at the Crichton Royal are included in that analysis. Lastly, among the young children and throughout all the age groups, as far as possible, care was taken to ensure that the distribution in terms of intelligence was uniform. The children from Gillespie's High School were looked upon as belonging to the top quarter, and the children in the other schools as representing the 50th and 75th percentile of intelligence. Again, in the other age groups, wherever possible about 20% of the people included are known to have University degrees or equivalent diplomas, thus ascertaining that approximately a quarter of the members of each age group belong to the top 20% of the population in terms of intelligence. It was however, impossible to ascertain the exact intelligence level among other subjects, but these are people such as artisans, shopkeepers, shop assistants, civil servants, secretaries, typists, workers, etc. representative of the general population.

In any study of a large population it is rather difficult to ensure that the sample is truly representative. As it is impossible to select the subjects completely, without bias, one has to accept a slight degree of bias as inevitable.

As far as the intelligence level and social background of the subjects participating in this study are concerned, a deliberate bias was introduced as has already been described in the previous paragraphs.

Another factor affecting the randomness of a sample can be introduced if it is assumed that people who come voluntarily for testing are representative of a given population. Chapanis (1950) in his study of a population tested at the Baltimore Exhibition, made such an assumption - those who came to the testing booths were looked upon as a good cross-section of the people who came to see the Exhibition. This assumption is suspect.

It is the experience of those who carry out routine colour vision testing, that there is a certain resentment on the part of many colour defectives when asked to do such testing, and it is often difficult to obtain the co-operation of people with only minor colour defects. Thus, it seems probable that these very young people would refuse to come forward for testing on a voluntary basis.

When it was decided to test a given group in this study, all the members of this group were persuaded to participate. For example, this was true of the school populations. This, of course, often entailed long delays in getting the whole group tested - in the case of a class at Motherwell Central School, it took about 12 Saturdays to test 40 boys. This prolonged testing time and was mainly due to the 'constant forgetting' on the part of 3 boys in the group to come for testing on the appointed Saturdays. When, finally, after much bribing they did come, it was found that two of these 3 boys were deuteranomolous subjects, and one had quite enlarged matching ranges in all 3 anomaloscope equations.

In those cases where passers-by were asked to participate in the experiments, only on four occasions was there a refusal to come for testing. This constituted a rather low incidence of refusals for this class of subjects and it is rather difficult to explain why this should have been so. The explanation might lie in the fact that many of the passers-by, who consented to come for testing were mistaken about the nature of the experiment. At the end of the testing session, they would say - "I thought that you asked me to participate in experiments on 'colour television'".

(c) Age Distribution

To help in the evaluation of results, an age order has been imposed on the data. It was rather difficult to decide on an age span as any choice is necessarily an arbitrary one. However, it was thought that it would be better to arrange the data in age groups in order to have the 10th, 20th, 30th etc.

years in the middle of the distribution for each age group. Each decade therefore starts in the middle of the previous decade - thus, the age groups are 16 to 25, 26 to 35, 36 to 45, etc. so that the 20, 30, 40 and 50 points are in the middle of the distribution for each group. Each age group covers 10 years except for the two youngest ones, where the span is of five years (from 5 to 10 and from 11 to 15). The table below shows the distribution of the mean age for the various groups for both sexes. Note that the age distribution of the last group (No. 8) extends from 65 plus to 90 years of age.

Table No. 1 showing mean ages for the eight sub-groups

	Age Group	Male	Mean Age Female	Total
1.	5 - 10	8.59	8.36	8.47
2.	11 - 15	12.95	14.13	13.53
3.	16 - 25	21.24	21.36	21.30
4.	26 - 35	29.52	29.84	29.68
5.	36 - 45	41.20	40.00	40.60
6.	46 - 55	51.12	50.92	51.02
7.	56 - 65	59.10	60.43	59.80
8.	65 +	73.33	74.32	73.80

3 : 3 PSEUDO-ISOCHROMATIC PLATES

Now we shall deal with the results obtained in our study of age and colour vision. Perhaps, a brief recapitulation of the evidence for changes in colour vision from these tests is advisable. In the first part of this thesis it was noted that where pseudo-isochromatic plates were used, these were of the type that tested only red and green colour discrimination. Secondly, though occasionally a battery of such tests was used (for example by Chapanis) the test most commonly used was the Ishihara. Thirdly, the losses that were expected in the aged were thought to be similar to those found in congenital colour defectives. Fourthly, some studies showed that in subjects of 60 years of age and over there was a higher incidence of people with great losses in colour discrimination, judging by their high scores (i.e. many misreadings on the Ishihara type of plates). Lastly, other results obtained from similar researches, especially those of Chapanis, were negative. What then is the explanation for these negative results ? This was partly due to the fact that older people were expected to make a large number of mistakes and secondly, to the fact that the analysis of results was always in terms of rather crude criteria.

It will be shown that, using both the Dvorine and Ishihara tests, changes which can only be attributed to age can be detected from the type of responses given, and that such changes are specific to a given age group. Colorimetric and photometric analyses will also be made to show why only fine analysis of the responses will lead to any indication of those changes. In the light of what has been found it will be obvious that to expect gross losses would <u>not</u>

be reasonable.

(a) Ishihara

I. Description of Test - The first of the tests to be discussed in detail in the Ishihara, more specifically the eleventh edition of this test. It is important to mention the edition because there are differences in printing and in general set-up of the test. Some earlier editions have only twelve plates whereas the later editions consist of 24 plates with digital and also non-digital plates for the use of children and illiterates. In this research the second part of the test has not been used at all, therefore all data given refer only to the first part of the set. Plates intended for testing the colour blind consist of six series each containing four plates.

In describing the series the same denomination of the various series as given by Le Grand, H. Hardy, in ~~his~~ their 1947 paper, will be used.

Series 1 is called a transformation pattern series the digits being composed in blue-red and yellow-red discs varying from 1 to 5 mm in diameter. This is seen by the colour normal, on a background of yellow-green and blue-green discs of similar sizes. Discs of both digit and background are intended to differ in saturation and luminosity.

In Series 2 (also called a transformation series) the colours of background and digit are reversed i.e. the blue-green and yellow-green discs of the digit appear on a background of yellow-red and blue-red discs.

Then we have Series 3 and 4. These are called the vanishing digit series and were originally conceived by Stilling. In series 3 the digit is formed of yellow-red discs on a background of yellow-green and blue, and in series 4 there are blue-green discs on a background of blue-red and yellow. These digits are plain to those with normal colour vision but are rarely seen by colour defectives.

Series 5 contains 'hidden digits' which are supposedly seen only by colour defectives but may also be visible to colour normal subjects.

Lastly we have Series 6 which consists of plates with two digits - one composed of red and one of red-purple discs appearing on a background of grey discs of differing brightness. Protanopes are supposed to see only the red-purple digit, while deuteranopes see only the red one.

II. Results obtained using the Ishihara -

(i) The analysis of the performance of subjects on a dichotomous colour vision test for inherited colour defects may be attempted in many ways. Individual performances on each plate can be analysed according to age-group. The type of misreading peculiar to a given age could be another form of classification. However, because of the nature of the test, it would perhaps be justifiable to start with the analysis on the one hand of the frequency of those with 'perfect' reading and on the other hand the incidence of those who make gross mistakes on their tests. By 'perfect' reading is meant readings where all the plates were read without any mistakes whatsoever, or if there were corrections they were made spontaneously, without probing by the tester. Of those with 'gross' misreadings, no distinction is made between those with genetically determined defects and those in whom there appears to be no genetic basis for any misreading.

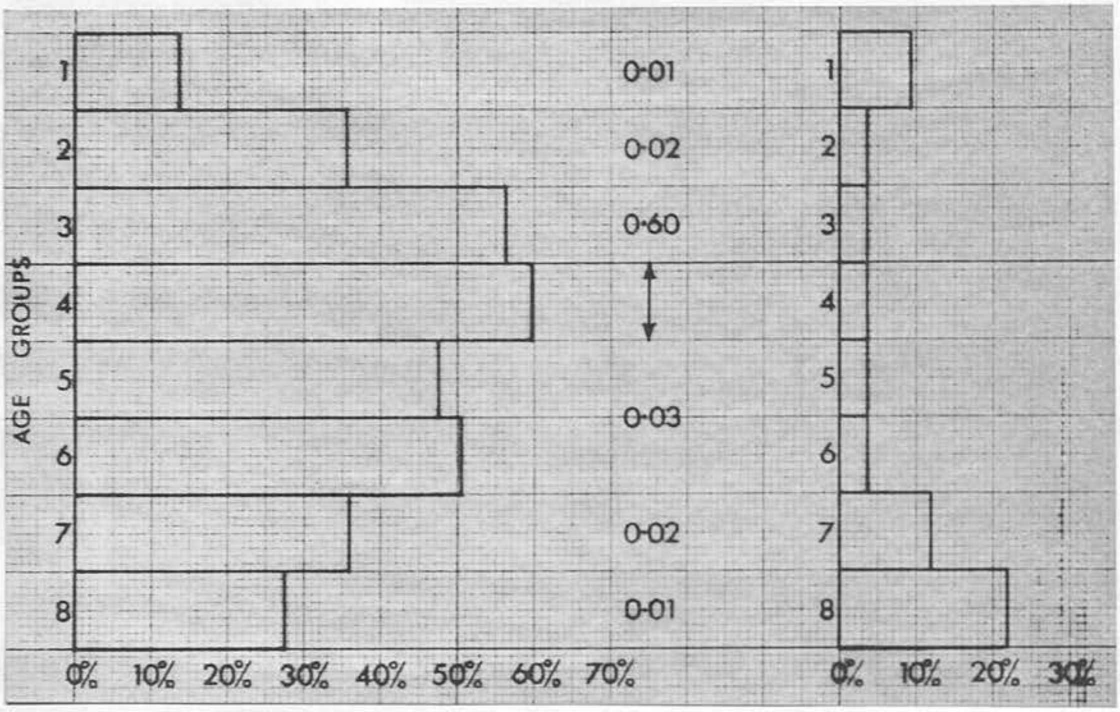


Fig. 1 - Distribution of perfect readings (left) and high error scores (right) for the eight age-groups tested using the Ishihara.

The diagram facing this page shows two histograms. On the left is a frequency distribution for those with perfect readings and on the right one for those who made seven or more mistakes.

Let us look first at the histogram on the left. This shows a gradual build-up of these with no mistakes in each group starting with the 5 - 10 age group till in the 25 - 36 group the frequency of those with no mistakes is highest. After this, there is a gradual decrease in the incidence of those who read all the Ishihara plates correctly. At the right hand side of this histogram there is a column giving the significant levels of differences between the percentages of 'perfect' readers for the age-group with the highest percentage, and each of the other age groups. An inter-group percentage differences of 18 is significant at the 0.05 level, while differences of about 25 per cent or over are significant at the 0.01 level.

The right hand histogram gives the frequency of the incidence of people with 7 or more errors for each age-group. There is a high incidence of such persons in the youngest age group, and also in the last two age groups (that is from 56 onwards). The differences in the percentage incidence of the two oldest age groups are significant at the 0.01 level.

(ii) Subsequent diagrams show the <u>distribution of percentage misreadings per plate for each of the age groups</u>.

When the eight age groups are scrutinised, the configuration of percentage misreadings for the total group can be seen. For example, in the 5 to 10 age group, most plates are misread on an average in about 4% cases while there are a few outstanding plates where the percentage is significantly

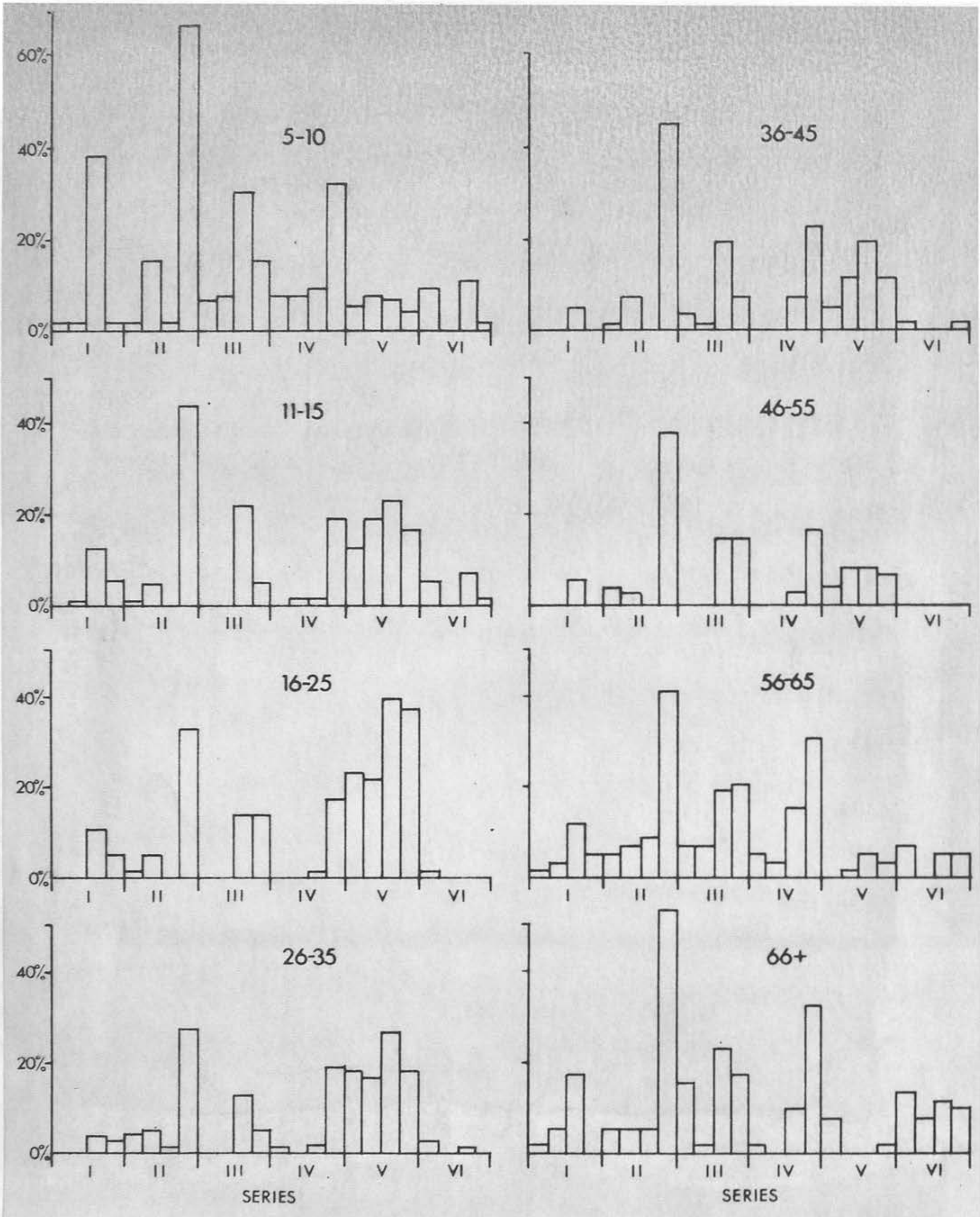


Fig. 2 - Distributions of percentage misreadings per plate for each age-group.

higher, (i.e. plates reading 29, 74, 97 and 73 all show over 30% misreadings). Scrutiny of the other age groups shows comparative findings. In the pocket provided on the opposite page, there is a transparency, giving the configuration of an average frequency distribution of misreadings for the whole population. By placing this transparency over any given age group it is possible to find by comparison how much over or under this 'average' any given age configuration is.

Another interesting picture emerges if we simply compare one series in each of the age groups. Take for example series 6 designed for the detection of dichromats. The incidence of misreadings found in the 5-10 and 11-16 age group, is in the order of 5 to 6 per cent for each plate in this series. In all subsequent groups except the last two the incidence of misreading is almost nil, while in the last age group it rises to about 11%.

The incidence of misreadings for series 5 (in this case reading the hidden digits correctly) presents another interesting distribution when compared from age group to age group. The highest incidence of these readings is in the 3rd age group (totalling about 40 per cent for all 4 plates), while there are hardly any readings of this series in the last age group. This point will be fully discussed in a separate paragraph dealing only with the performance on the Hidden Digit series.

(iii) The results may also be analysed in terms of the <u>number of misreadings made per age population.</u> The diagram showing the average number of misreadings per person shows us that this varies from 3 per person in the 5 to 10 age group to 1.2 per person for the 46 to 55 age group and then again increases to 3 mistakes per person in the two last age groups.

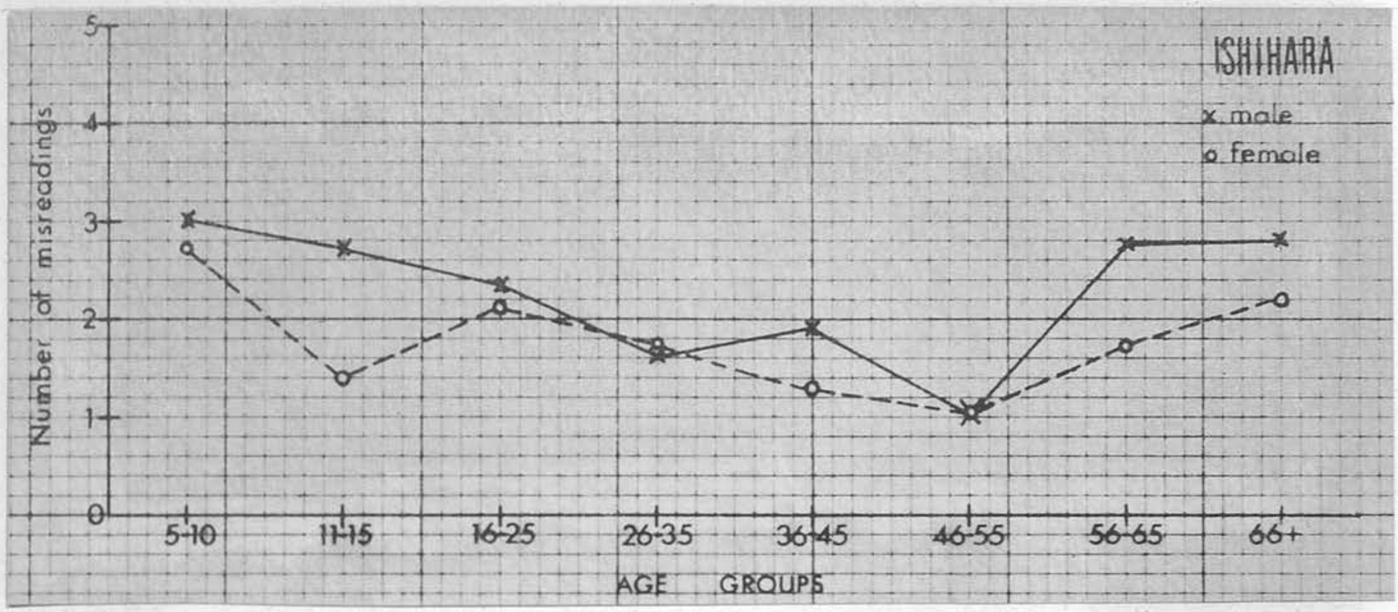


Fig. 3

(iv) The same thing happens when we take the total percentages of misreadings per age group. Here, among the very young, approximately 13% of all the plates in the test are misread and this percentage gradually decreases till we come to the 46 to 55 age group where only 5% of all plates are misread. Then, from 55 years onwards, there is again an increased percentage of misreadings.

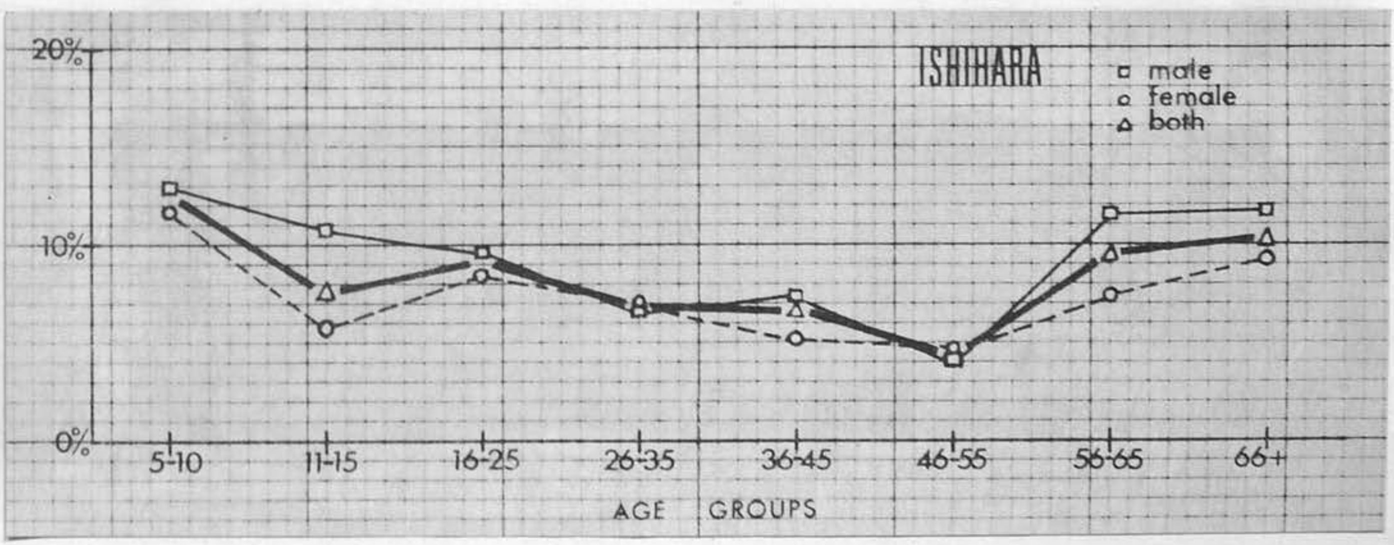


Fig. 4

ISHIHARA PLATES

Total frequency of misreadings per plate - (total population)

Series	No.	Correct Reading	Both Sexes Frequency	Both Sexes %	Female Frequency	Male Frequency
	2	8	3	.587	2	1
1	3	6	7	1.37	4	3
	4	29	60	11.74	30	30
	5	57	14	2.74	7	7
	6	5	15	2.94	4	11
2	7	3	36	7.45	16	20
	8	15	10	1.96	5	5
	9	74	215	42.07	96	114
	10	2	17	3.33	7	10
3	11	6	11	2.15	4	7
	12	97	90	17.61	46	44
	13	45	63	12.33	33	30
	14	5	10	1.96	6	4
4	15	7	7	1.37	3	4
	16	16	27	5.28	10	17
	17	73	115	22.5	61	54
	18	0-5	50	9.78	22	28
5	19	0-2	56	10.96	20	36
	20	0-45	79	15.46	25	54
	21	0-73	63	12.33	20	43
	22	26	25	4.89	10	15
6	23	42	6	1.17	2	4
	24	35	22	4.31	10	12
	25	96	13	2.54	5	8
		Total	1014	Average 8.26 %	448	566
					44.18% or	55.82%

Males 251) Total 511
Females 260)

Possible No. of mistakes 12,264
Average No. of misreadings (Males 2.25) (Females 1.72) per person

In this type of analysis of the Ishihara, there is a gradual improvement with age up till 50 judging by the decrease in the number of mistakes per person. This is consistent with all the previous results published by Chapanis and others, who either noted such trends or found no difference in the number of mistakes made per person, over the 20 to 50 age spans.

(v) There are interesting differences between the two sexes in the incidence of the average number of misreadings per person per age group. Females make fewer mistakes in the 11 - 15 age group and also in the last age group. In all other age groups differences, where they exist, are very small. The same comment applied to the diagram showing total percentage of misreadings per age group.

Because of the interest of previous research workers in finding if any differences in performance on the Ishihara existed between males and females, a further analysis of the frequency of misreadings is made. A table showing the total frequency of misreadings per plate for the whole population is given. The plates are given their serial numbers for the sake of clarity, and dealt with in blocks of four, as in the six series of the test. The correct number for the normal observer, is quoted next. Then the frequency of misreadings is given for each plate and following this are listed the percentages that these numbers represent out of the total number of misreadings of all the plates taken together. Lastly, the two columns on the far right give the number of misreadings of females and males for each individual plate.

The total frequency of misreadings for females is 448 and for males 566. This represents a percentage misreading of 44.18% and 55.82% respectively, for females and males. The 11.6% difference is significant at the 0.01 level.

However, if we look at the table, it will be noticed that discrepancies of any magnitude between the sexes apply only to series 2, 5 and 6. If we exclude these three series from the total analysis it is found that the difference (230 females, 255 males) is no longer significant.

Using the X^2 technique to test the significance of the difference between the misreadings for males and females for the 3 series we obtain the following results :-

	Frequency of misreading. Female	Male	X^2	P
Series 2	121	150	3.114	0.08
Series 5	97	161	15.87	< 0.001
Series 6	27	39	2.2	< 0.15
Total test	448	566	13.73	< 0.001

Table No. 1 Chi-square for the frequency of misreadings by males and females.

Thus only differences obtained on series 5 are significant at the 0.01 level, while in series 2 a significance nearing 0·05 is obtained. The greatest load comes then from series 5. As has already been mentioned, this series is 'misread' most frequently in the 15 to 25 age group - therefore it is fair to assume that the differences existing between males and females when the total number of misreadings are compared comes mainly from the differences in misreadings of series 5 and chiefly from the 16 to 25 age group.

This point will be more fully elaborated when series 5 is discussed, and in section (d) where reasons for the difference in colour discrimination due to age will be discussed.

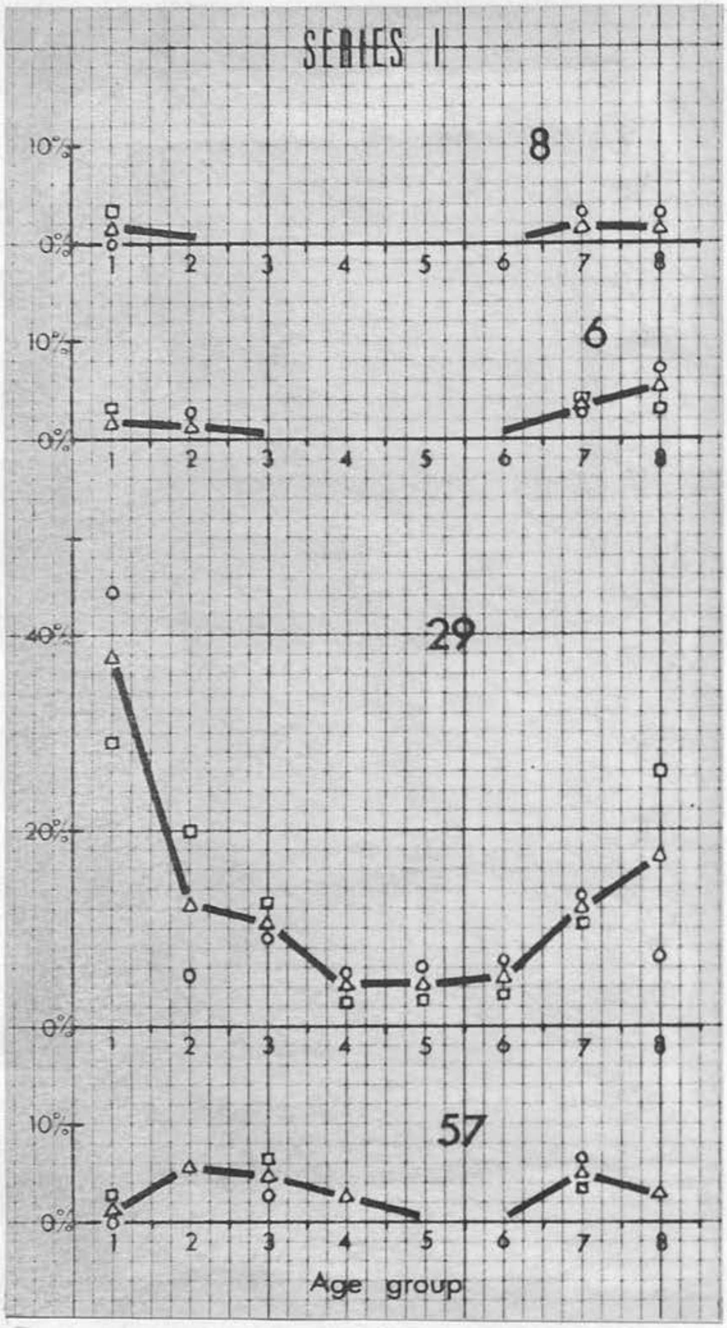


Fig. 5 - Frequency of misreadings per individual plate - series 1.

(vi) Finally, there is a set of diagrams giving the frequency of misreadings per individual plate for each of the separate series. These diagrams show that there are some plates that are hardly affected by age variations, while others are extremely sensitive to them. For the purpose of this, and the previous analyses the scores of all major defectives have been excluded.

In the first series, plates with one digit show only a slight incidence of misreadings both in the first age group and in the last two age groups. Of the two double digit plates in this series, the one reading 29 is most frequently misread by the very young and to a lesser extent by the older people, but it is also misread by 2 to 3 per cent of all other age groups.

The most frequently given misreading for this plate is 20.

Table No. 2 shows the incidence in percentage of such reading for plate No. 4 throughout the various age groups.

Table No. 2 Series 1 - No. 2 reading '29'

Age Groups	1	2	3	4	5	6	7	8
Read as 20	34.8	9.5	3	2.5	0.5	3	3	8

With the plate reading 57 (i.e. plate No. 5) the highest incidence of misreadings is in the third age group, whereas when we come to the seventh and eighth age groups, misreadings decrease significantly.

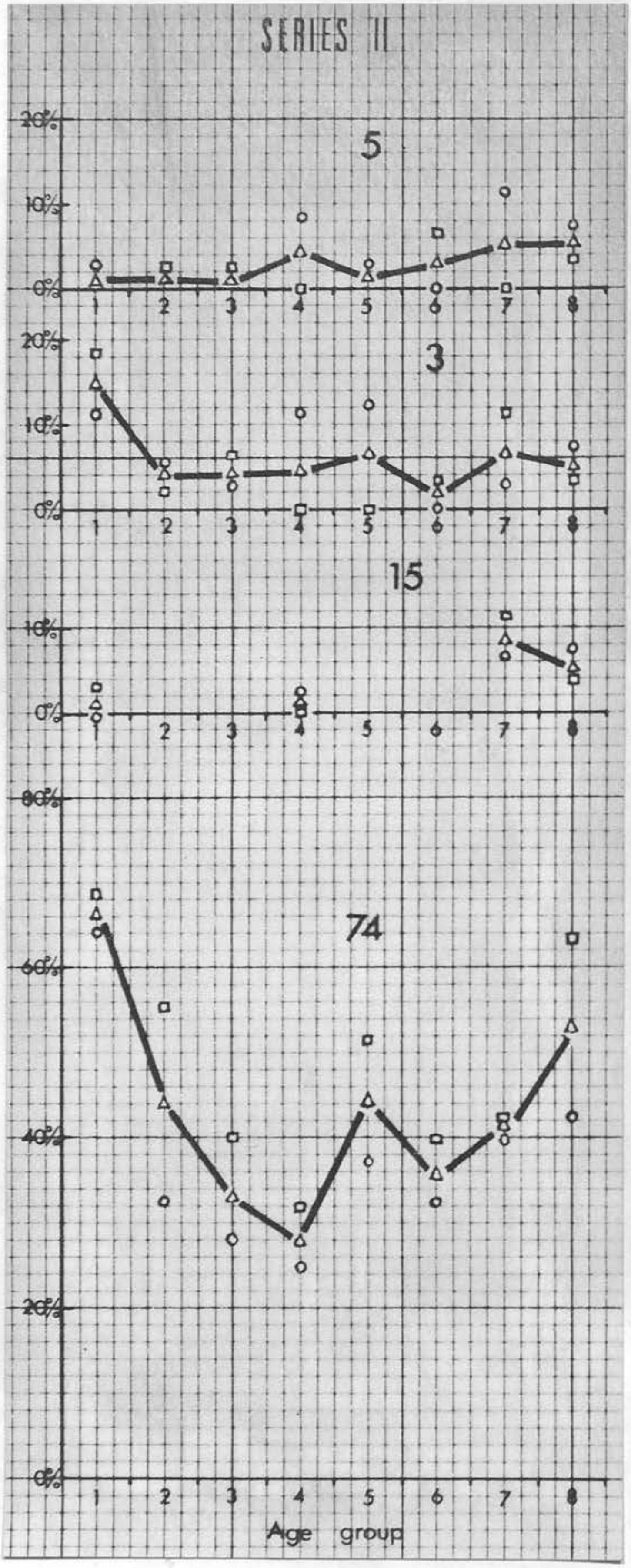


Fig. 6 - Frequency of misreadings per individual plate - series 2.

In series 2 again we have two plates with a single digit and two plates with double digits. These first two plates in this series are noticeably less affected by age variation than the last two, but we find that the general incidence of misreadings is much higher here than it was in the first series. Of the double digit plates in this series the one that is least affected by age variation is plate reading No. 15, while plate reading No. 74 has the highest misreadings per age group and also has the highest overall misreadings for all the plates in the Ishihara test. This plate will be discussed more fully later in the section dealing with colorimetric analysis, and an explanation for this given.

The following is a table of the incidence of 'correct' reading of plate No. 9 for the various age-groups. Percentages are also given for the most common type of misreading (i.e. 71) of this plate

Series 2 - No. 9 reading 74

Age Groups	1	2	3	4	5	6	7	8
Read as 74	33.3	56	66	72.4	55.5	63.9	58.6	47.1
Read as 71	50	27	16.9	16.6	23.8	19.6	20	28

Note that the plate is correctly read most often in the age groups 3, 4, 5, 6 (that is between the ages of 16 and 55 years), whereas the incidence of reading it as figure 71 is least frequent for these age groups. However, half of the subjects in the 5 - 10 age group read it as 71, and only 33.3% read it as 74. On the principle of most frequent response the 'correct' reading for this age group should then be '71' and not 74.

In series 3 the pattern so far described in the previous two series

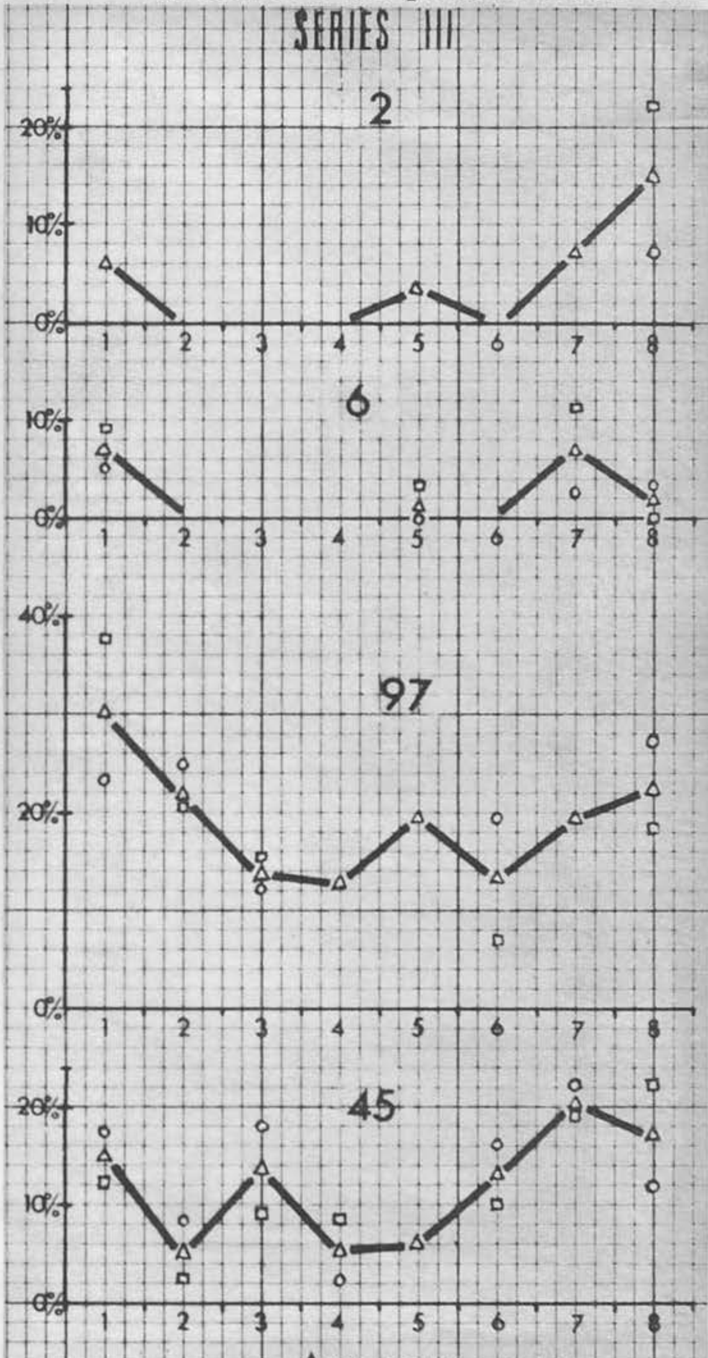


Fig. 7 - Frequency of misreadings per individual plate - series 3.

repeats itself - the double digit plates have the highest overall number of misreadings, though the single digit plates show a much higher incidence of misreadings for both the young and the old subjects than for the middle groups. In the last two plates of this series, the commonest type of mistake is a misreading of either the first or second number of the double digit figure. This applies more frequently for the first four age groups, though in the older subjects there is a relatively high incidence of those who instead of reading 97 read 87 for the third plate in this series. In this plate then, the second digit is less frequently misread, while '9' (the first digit) is misread as follows :-

Age Groups	1	2	3	4	5	
Read as 87	25	20	12	13	11	%

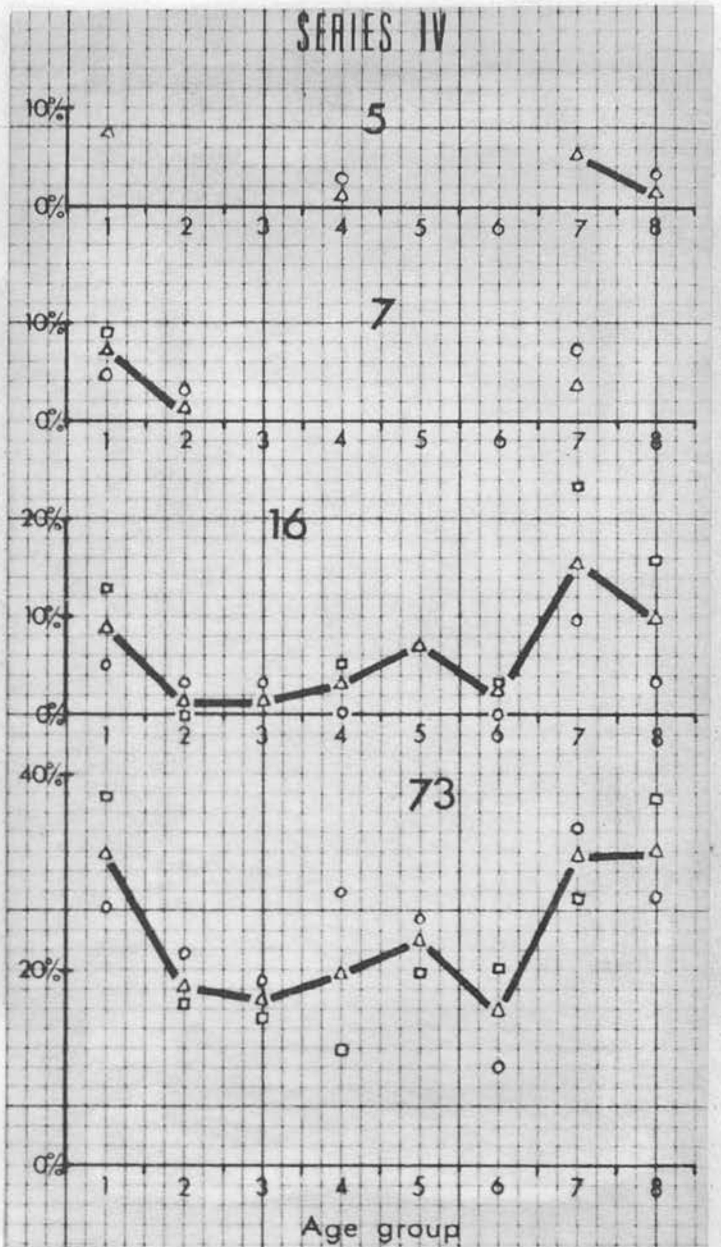


Fig. 8 - Frequency of misreadings per individual plate - series 4.

In subsequent series, i.e. series 4 again the double numbered figures show a higher incidence of mis-readings for any population, but there is also a general increase in the two oldest age groups, whereas the two single digit plates show hardly any effect due to age. The last plate in this series reading '73' is the second most frequently mis-read plate in the Ishihara, about a quarter of all misreadings come from this plate.

The common misreadings are 23 or read as one digit figure only, be it 3 or 7.

Reading this plate '23' occurs more often in the 4, 5, 6 age groups while reading only 3 is confined to the two last age groups. On the other hand only in the 3 youngest age groups do we find the misreadings of the second digit in this plate. Here is a table summarising the position.

Series 4 - No. 13 reading 73

Age Groups	1	2	3	4	5	6	7	8
Read as 73	68	80	83	80.5	77.7	85.2	68.9	67.9
Read as 23	-	1.4	3.0	7.0	15.8	8.5	3.5	10
Read as 3							8.0	10
Read as 7 only	27	15	12					

Thus in this plate read as 73 by the normal observer, misreading of the first digit is the most common mistake.

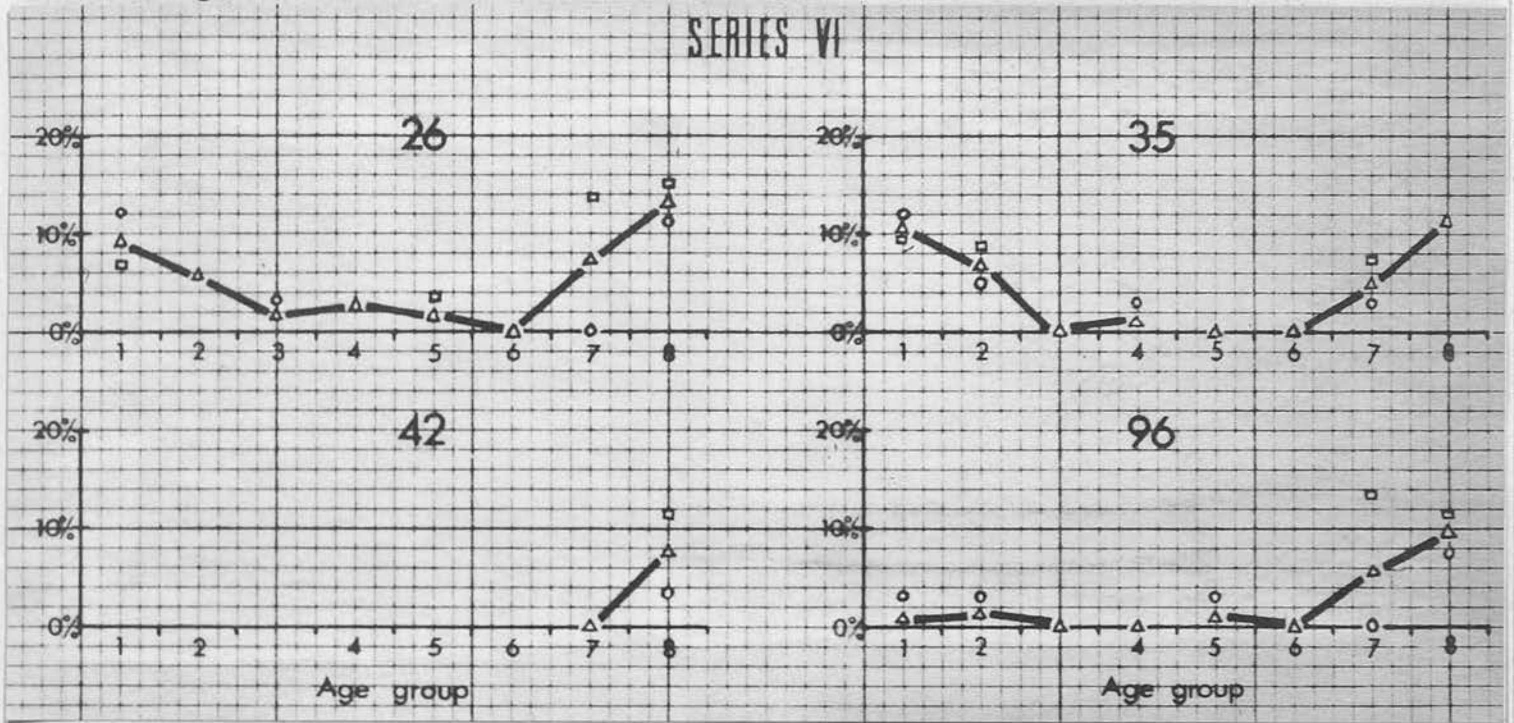


Fig. 9 - Frequency of misreadings per individual plate - series 6.

Before the discussion of series 5, we shall look at series 6 dealing with the special plates that differentiate between the two types of dichromats. The plate reading 42 has the smallest incidence of misreadings of almost all the plates in the test. The other three plates are misread both by young and old, the incidence for the last age group being in the order of 10 per cent for these 3 plates.

(vii) Of special interest are the four plates in series 5 - the so-called hidden digits. Most research workers say that, from the diagnostic point of view, they are useless except in that they have a psychological effect on the colour defective person who is able to read some of them. Some others, such as Pickford and Chapanis, say that a lot of normal people can read them as well. From this research a rather interesting and perhaps surprising picture has emerged.

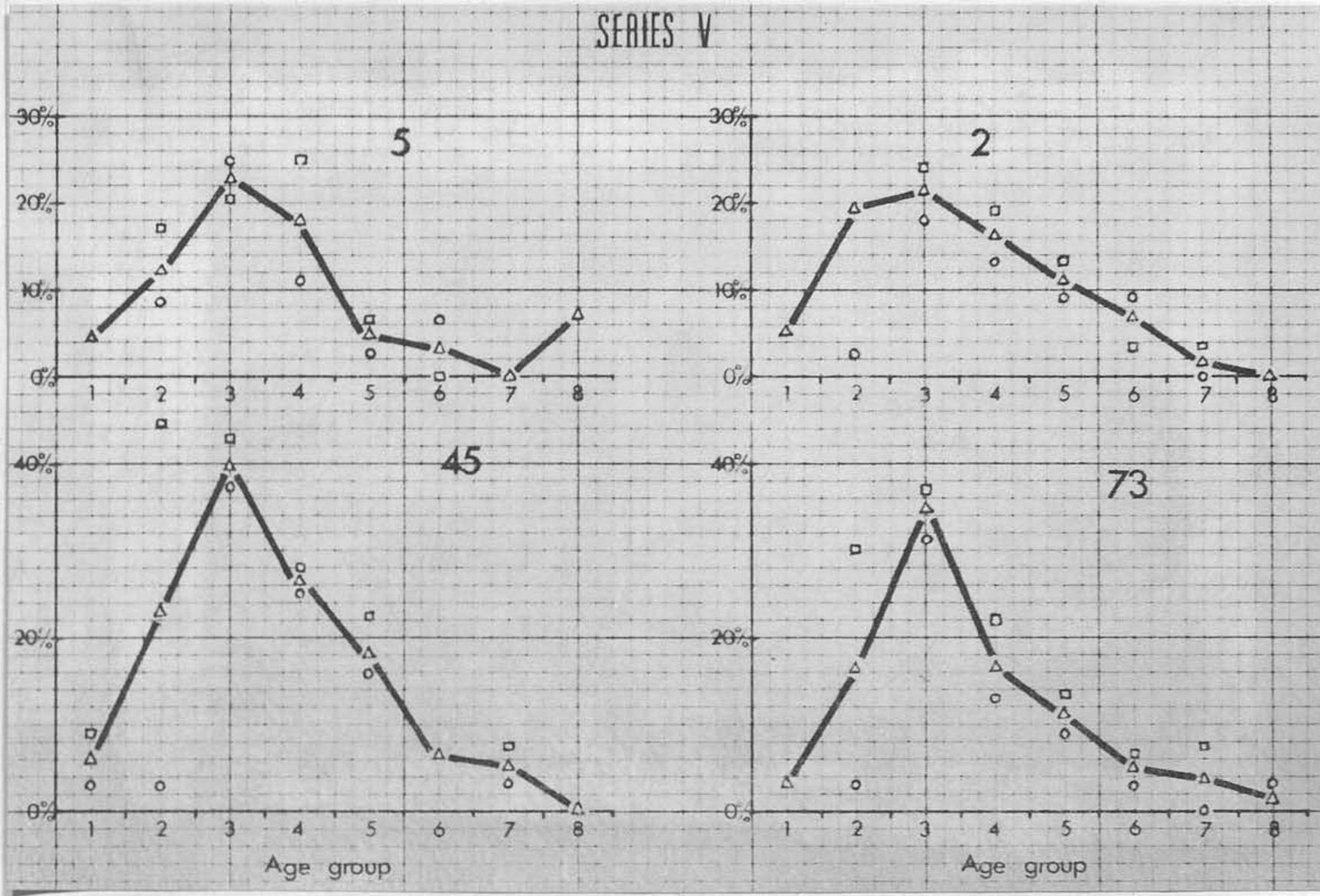


Fig. 10 - Frequency of misreadings per individual plate - series 5.

Whereas most of the curves for the other plates are either 'flat' for the whole age distribution, or show a 'u' type of distribution, the incidence of correct readings (i.e. in the same way as would the colour vision defectives) gives us a reverse type of curve, an upside down 'v' with the peak in the third age group (i.e. between 16 and 25 years of age). There is a degree of variation between these four plates, i.e. the third one No. 45 shows the highest incidence of perfect readings, that is, about 40% in this third age group read it correctly. Notice how steeply the incidence drops from this third group till in the last age group the readings are generally nil. Old people over 50 or 60 seldom read these plates.

As has been already mentioned, the sex differencesin being able to read these hidden digits are quite marked for this series. This difference is the largest for all the series in this test. The diagrams showing the incidence for the various age groups show, that these differences between the sexes are greatest in age groups 2 and 3.

The small differences for other age groups are not significant at all, thus we may conclude that the 'age variable' is an important factor, in determining when such digits will be read by subjects with normal colour discrimination.

However, though the hidden digit recognition is dependent on age, it is also dependent on the total number of mistakes a given person makes on the Ishihara. This is demonstrated in the next table, which gives the incidence of reading the hidden digits among the subjects of 1st Ordinary Class of Psychology tested in 1959 on the Ishihara.

Ishihara - Hidden Digits

Student Population 1st Ord. Psych. 1959 - N = 297

Total Number of Mistakes	No. read of H. D.	/	Total No.	Average No. of reading H. D. per person
0	80	/	105	0.83
1	72	/	63	1.14
2	65	/	54	1.20
3	46	/	33	1.39
4	31	/	18	1.72
5	7	/	2	3.5
6	13	/	6	2.16
8 - 11	13	/	6	2.16
18 +	36	/	10	3.6

The table is constructed in such a way that it gives us a column of misreadings, a column of fractions (i.e. the number of misreadings per group of people over the total group) and finally a third column showing the average number of hidden digits read per person in that group. The grouping here goes according to the number of mistakes made (on 20 plates) with results for no mistakes; 1, 2, 3, 4, 5, 6, 8 - 11, and 18 plus. In the next column we find that, among those who made no mistakes in the 20 plates, 87 out of 105 read one of the four hidden digits thus giving an average number of correct readings of 0.83 per person. Taking the fourth group, (those who made three mistakes on the 20 plates) we have 33 people who read altogether 46 plates in the hidden digit series correctly, giving an average number of hidden digits of 1.39 per person in this group. Then let us take the last group - those who made 18 mistakes or more (i.e. dichromats). There are 10 subjects in this group and these 10 made 36 correct readings giving us an average incidence of 3.6 of the hidden digits read per person. Thus reading the digits correctly is not only a function of age but is also a function of the degree of anomaly a person possesses. The higher the misreadings of the 20 plates in the Ishihara, the higher is the number of correctly recognised numbers on the 4 hidden digits also.

(viii) The last comparison of data on the Ishihara is of the incidence of those making 1, 2, 3, 4, 5, 6, 7, mistakes for each of four age groups.

For the purposes of this comparison these four sub groups will be (a) from 5 - 15 (b) 16 - 35 years (c) 36 - 55 years and (d) 56 plus. Each of the sub groups contains around 130 subjects as can be seen from the following diagram showing the age and incidence of error comparison for the whole population.

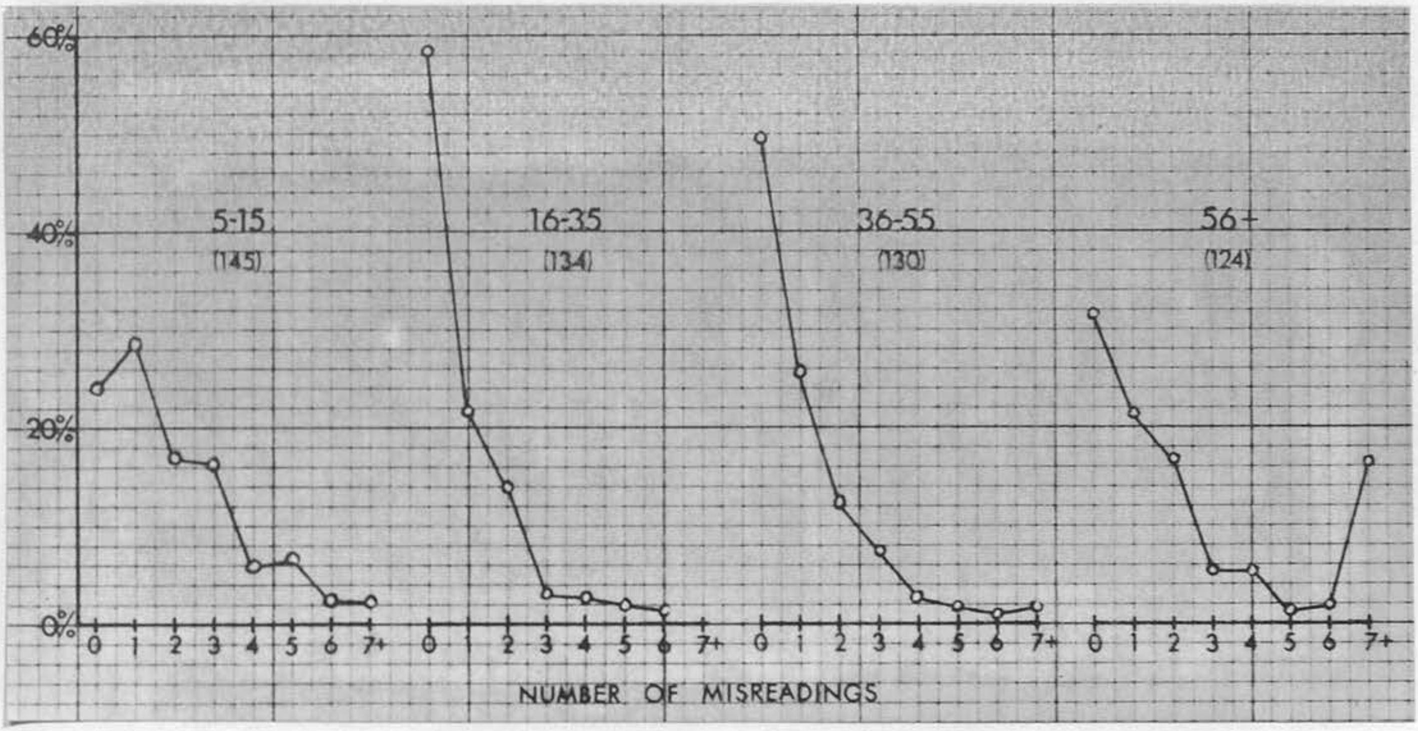


Fig. 11 - Frequency of misreadings per age-group.

The percentage incidence of errors is measured along the vertical axis, and the number of misreadings starting from 0, 1, 2 etc. is plotted along the horizontal axis, (the last position [7] on the horizontal axis includes all those who made 7 or more mistakes). All major defectives have been excluded.

Only in the 16 - 55 age groups do we see a reversed J type of frequency curve. For the young group, though there is a trend towards such a curve, it is very irregular in shape.

In another respect this frequency distribution (i.e. for the youngest group) is different from all other age groups, in that the incidence of those who make no misreadings, is less than the incidence of those making one mistake. This is not found in any other frequency distribution quoted by other researchers (except Crawford).

In the last age-group (56 +), the incidence of those who make 7 + mistakes is about 18 per cent, and there is an increase in the incidence of those who make 3 and 4 mistakes in comparison with the two previous age groups. Again the number of those with no misreadings is quite small in comparison with the middle groups.

A comparison, of the performance of the 'Age Population' with other homogeneous groups can be made, but care must be taken when making such comparisons, to note whether the sample to be compared included major defectives or not. Some research workers tested an unselected population and some confined themselves to non-defective subjects, therefore, comparisons will be made with two types of population -

(1) inclusive - where everyone tested has been included even major defectives.

(2) modified - where populations of 'normal' subjects only were tested.

The groups differ from each other in one further respect - that is in the way they were collected. These groups may be termed either as closed or open populations. By a 'closed' population we refer to one in which everyone in a given class, locality, or other homogeneous group has been tested, and by an 'open' population we refer to a population where some choice was made, and where only a 'sample' of a given class was tested.

Table No. 3

Student Population - 1st Ord. Psych. 1959 N. 297

DISTRIBUTION OF NUMBER OF MISREADINGS - ISHIHARA

	Males		Females		Total	
Number of Mistakes	No.	%	No.	%	No.	%
0	51	36	54	35.2	105	35.0
1	28	20	35	22.4		
2	18	12.8	36	23.0	117	39.4
3	18	12.8	15	10.0		
4	8	5.7	10	6.0	51	17.1
5	1	0.7	1	0.7		
6	3	2.0	3	2.0	8	2.6
8-11	5	3.5	1	0.7		
18+	9	6.0	1	0.7	16	5.4
	141	99.6	156	99.7	297	99.5

Table No. 3 on the opposite page giving the distribution of misreadings of the Ishihara is an example of a closed population where the whole of the first Ordinary Psychology Class was tested.

The distribution of misreadings for a closed population is quite different from the distribution for any other type of population because there are always more defectives in such a group than the normal 7 - 8% quoted for a "random population". This student population, for example, has a total percentage of 10.4 making eight or more misreadings.

Some of the samples discussed here include equal numbers of female and male subjects, while other samples have a preponderance usually of males.

Table 4 gives a brief picture of the type of sample to be compared.

Researcher	Year	INCLUSIVE Male	INCLUSIVE Female	INCLUSIVE Total	MODIFIED Male	MODIFIED Female	MODIFIED Total	Type	Degree of randomness
Belcher et al	1958	468	32	500			462	Student staff	open, random
Crawford	1955	34	43	76	24	41	65	1st y. Psych. class	selected
Lakowski	1958				251	260	511	Age groups	open, random
Lakowski	1959	141	156	297	127	154	281	1st y. Psych.	closed
Sloan	1956						100	not known	selected

Table No. 4

Let us begin by comparing the 'inclusive' populations, homogeneous in terms of intelligence and social standing - the three student populations of

Belcher, Lakowski and Crawford. The frequency polygons on diagram No. 11 summarise the situation. There are more subjects making no mistakes in the Imperial College population, than in the two other comparable groups. However, the Lakowski and Crawford populations are in close agreement about the incidence of those with 1, 2, or 3 misreadings.

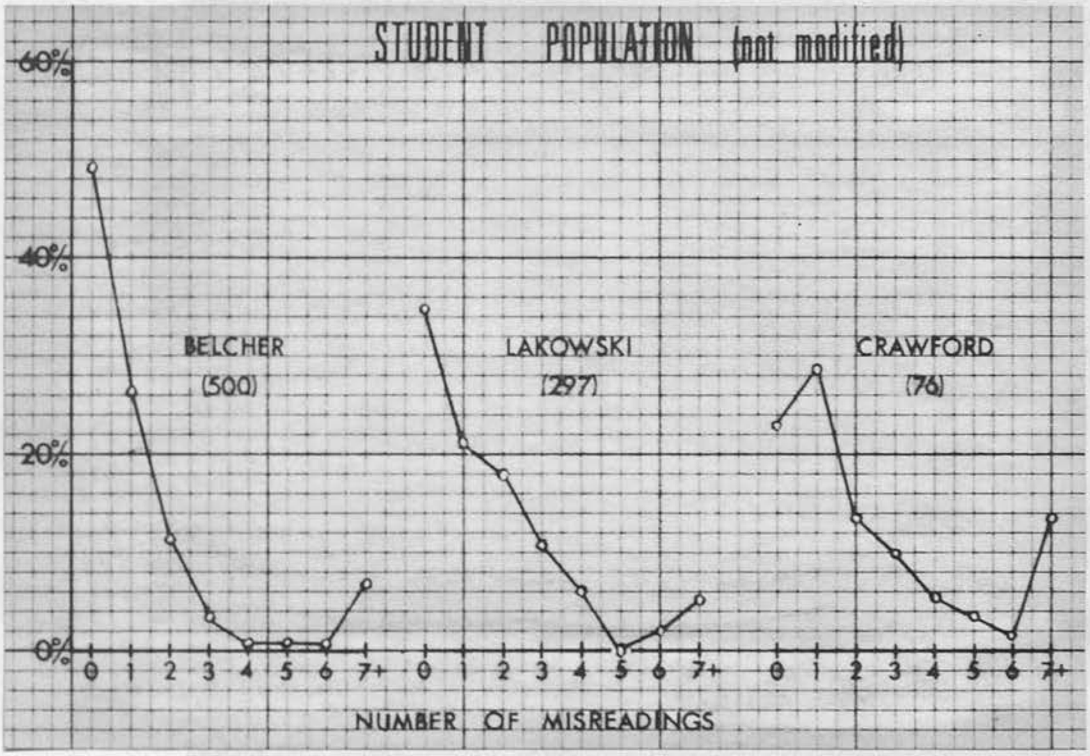


Fig. 12 - Frequency of misreadings - three student populations.

Next we shall compare the five modified populations, where major defectives were excluded. Here populations differ in terms of the type of subject participating in the tests, and also in the 'degree of randomness' of the population tested.

Diagram No. 13 summarises the results of 'modified' populations.

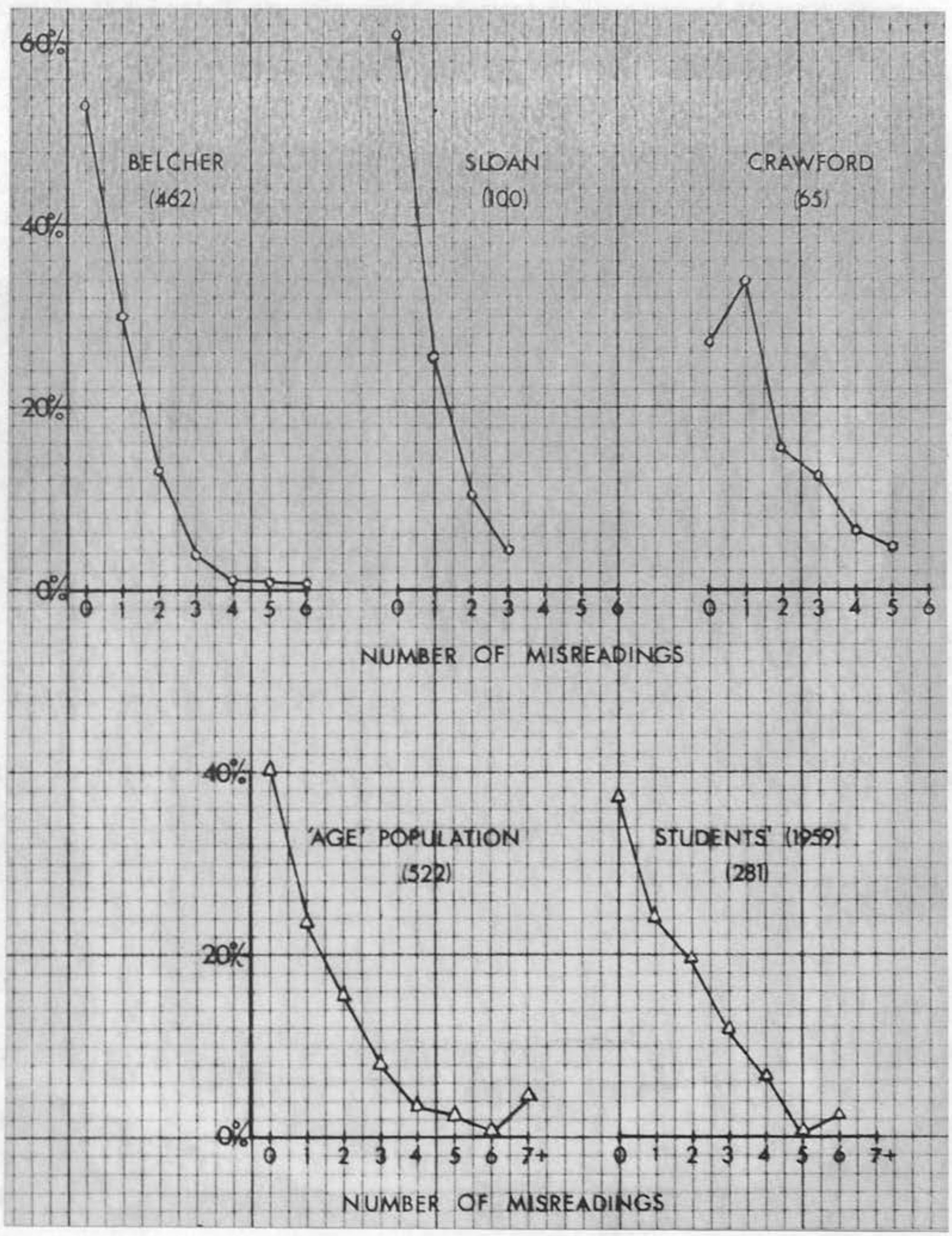


Fig. 13 - Frequency of misreadings - five 'modified' populations.

The first two populations of Belcher and Sloan, show a swiftly down-sweeping frequency curve of the inverted J type, whereas the other three populations have more gently sloping curves. If we compare the percentages of those making 3 mistakes in the last 3 groups we find that it is around 10 per cent, and note just 3% as shown in the first two populations. There are also more subjects making four, five etc. mistakes, in the latter 3 groups than we find in the populations of Belcher and Sloan.

Now referring back to the diagram showing the frequency curves for the four age groups (i.e. diagram No. 11), we notice that only the 16 - 35 age groups distribution of misreadings is similar to the distributions of Sloan, and Belcher. The curves for both the young and the old subjects differ markedly from such distributions.

III. <u>Summary of results</u> - The results obtained from the data on the Ishihara in this research do not confirm the conclusions of Chapanis, that this test is incapable of distinguishing age variations. They are more in accord with the work of Janouskova who showed that in the children and very old subjects a higher incidence of people with large numbers of misreadings is found than would be expected from the frequency of congenital colour defectives in the population in general. The test differentiates not only at the extremes of the age continuum, but also gives some indication of changes in perception occurring in the 20 to 50 age group, though it must be emphasised that the extent of the changes found is not very large. They could not, for example, be detected in the study of only one or a few subjects in each of the given age groups. There are trends and tendencies, which can be seen and measured only when comparisons are made between large group samples. In addition a highly analytical approach is essential - the responses of subjects must be analysed not in relation to the whole test but rather to individual items or series within the test. When such an approach is followed,then meaningful results can be obtained. It can for example, be said that some of the plates in the test are more sensitive in detecting age variations, while others are less so. Generally speaking the plates that give a clear cut distinction between normal and defective colour discrimination (as found in researches using the Ishihara as a Dichotomous test) are the very plates least effected by age variations. Plates with double digits seem to offer a better indication of age differences than single digit plates.

Of interest is the peculiar shape of the curve showing the frequency of readings of the hidden digits, where the plates seem to show a reverse effect of the influence of age on the readability of these plates (i. e. the younger subjects read them more often).

Pickford (1950) in an item analysis of the Ishihara test, concluded that if plates Nos. 10, 11, 14, 15, 18 and 23 were selected, an ideal dichotomous test would be available which could give clear-cut divisions between those with normal colour discrimination and those with defective colour discrimination. This choice of plates corresponds with the analysis of data in this research. In terms of age variation, the above are the plates that are least affected, thus confirming Pickford's choice of plates. A similar choice was suggested by Sloan (1945) and Collins (1937).

When we consider the results presented here, we have the interesting situation where depending on what is wanted from such a test, it is possible either to compose a revised test consisting of these six or seven plates alone, or to keep the test as it is, bearing in mind that the various misreadings obtained are typical neither of the normal person nor of the defective person. Nevertheless, they indicate certain minor changes which must be taking place in the visual system both of the very young and of the very old.

In the section dealing with photometric and colorimetric data, an explanation will be offered to show why some plates in this test are more frequently misread than others, by non-defective persons. It will be seen that the plates which are confused by the major defective subjects only are the very plates in which the colour of the figure and background lie furthest away from each other in the colour space, while the reverse is true of plates affected by the age variable.

(b) Dvorine

The second edition of Dvorine's pseudo-isochromatic plates used in this research consists of two sections, the first containing one demonstration plate and 14 plates showing numbers, and the second containing 7 plates plus a demonstration plate using lines instead of digits. (In this research the second part of the test has been omitted). Also included is a Nomenclature test which will be dealt with separately. This test is intended to be both a dichotomous as well as a qualitative test of red-green defects, that is, it is capable of selecting normal people from those with defective red-green vision and also contains plates which can differentiate protanopes from deuteranopes.

The test consists of seven series :

In series I, plates 2 and 3 are made up of red and green coloureddots reading numbers 67 and 38 for the normal observer.

Series II contains plates No. 4 and 5. These have a figure and background composed of brown and reddish dots reading Nos. 92 and 70.

The third series consists of plates Nos. 6 and 7. Here we have grey figures against a purplish background. These are the two plates that give a qualitative diagnosis - a subject with a protan type of deficiency identifies digits 9 and 2 in the two plates, whereas a deutan identifies a second digit in each plate i.e. 5 and 6.

The plates in series IV are Nos. 8 and 9 reading 2 and 74 respectively where the figure is orange against a yellow background.

The next series (plates No. 10 and 11) has a yellow figure on a yellow green ground, and the figures read 62 and 4 respectively.

In the VI series there are two plates reading 28 and 46 made up of greenish figures on an orangey background.

And lastly, (series VII) plates 14 and 15 read 7 and 39 respectively, and are composed of violet or purplish figures against a bluish background.

Of the 14 plates, only three are one digit figures - the others are all two digit numbers. The administration of these plates is similar to that of the Ishihara, that is they have to be read under a northern light illuminant at 30 ins. from the subject and a period of not more than 5 secs. is given for the identification of each plate. Perhaps it should be mentioned that the figures or configuration are less distinct than their counterparts in the Ishihara.

I. Results obtained using the Dvorine plates - The analysis of results follows the same pattern used for the Ishihara. Firstly, the incidence of perfect readings will be given followed by the incidence of gross misreadings. Secondly, the number of mistakes per age population and the various frequencies will be shown. Then the incidence of people who make 1, 2, 3, etc., mistakes in each age group will be analysed and finally the results of this research will be compared with those of other research workers who used the Dvorine on large populations.

Analysis of these results shows that

a) the frequencies of those who made no mistakes whatsoever and

b) those who made 7 or more mistakes

follows the same pattern for the Dvorine as it did in the Ishihara.

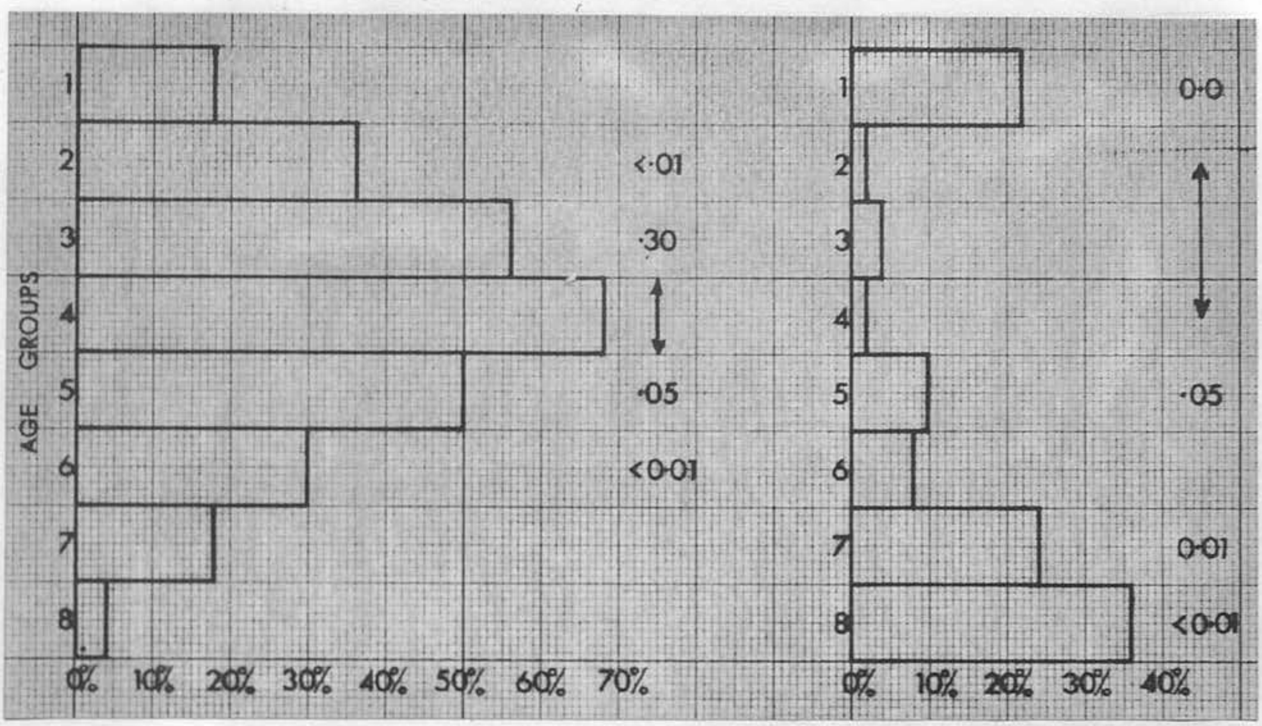


Fig. 14 - Distribution of perfect readings (left) and high error scores (right) for the eight age-groups tested using the Dvorine.

The first frequency histogram shows that there is a gradual build-up of people passing this test with no mistakes in each age group from the 5 to 10 year olds till we come to the 25 to 36 age group where the frequency is the highest. From there on, there is a gradual decline in the numbers reading the Dvorine plates correctly.

The second histogram giving the incidence of those with 7 or more mistakes shows that in the first age group there is a greater incidence of those making many mistakes than in the middle groups and this is again true of the last two age groups, where the incidence is as high as 24 and 36 per cent. This percentage for the last two age groups is much higher than was found using the Ishihara plates (where it was 12 and 22 per cent respectively). Already in the two age-groups preceding these (i. e. 36 - 55) there is an increase in the number of those with seven or more misreadings. The scores of major red-green defectives (as diagnosed by the anomaloscope) have been excluded from these frequency distributions. At the right

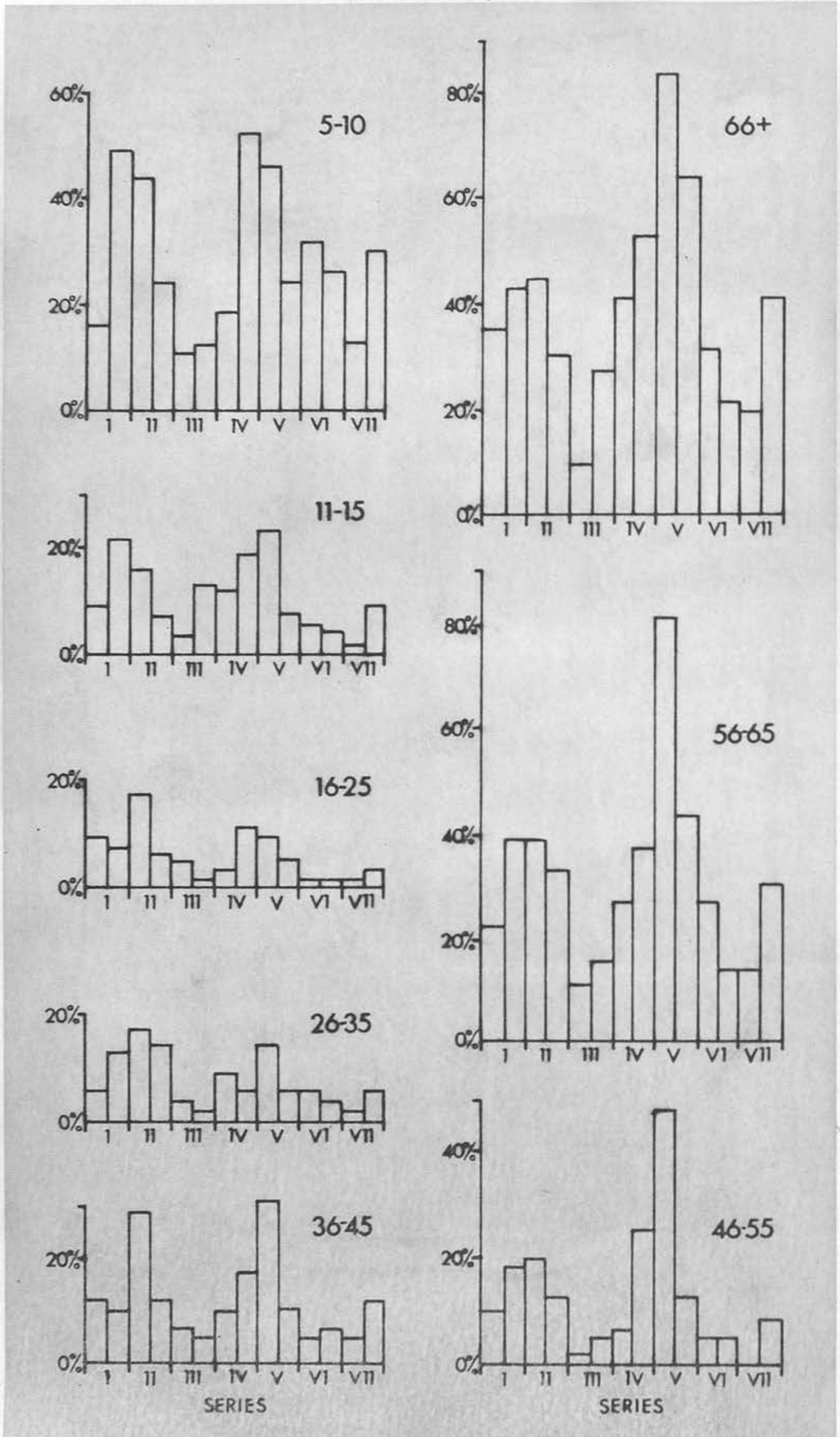


Fig. 15 - Distributions of percentage misreadings per plate for each age-group.

hand side of these histograms there are columns giving the significant level of differences between the age groups with the lowest and those with the highest incidences either of perfect readings or of misreadings.

Subsequent diagrams show the distribution of percentage misreadings per plate for each age group. From these diagrams we notice that there are two peaks of misreadings around plates 3 and 4, and 9 and 10 in all theage groups, giving a picture approaching a bi-modal distribution of misreadings. This distribution is least pronounced in the age groups from 16 and 35 and is most pronounced in the 65 plus age group. Plates 5 and 6 intended to detect dichromats of the red-green type are least affected by this age variable, and thusproduce the dividing line between these two peaks. In the pocket provided on the opposite page there is a transparency, giving the configuration of an average frequency distribution of misreadings for the whole population. By placing this transparency over any given group it is possible to compare it with the whole group and to see how much over or under this 'average' a given age configuration is.

Analysing the results in terms of <u>number of mistakes made per age population per person, an average of the number of misreadings is obtained.</u> It is found that in the first age group (5 to 10) the average is 4 mistakes; which is the same as was found for the 7th age group, while in the last group the average is as much as 5 mistakes per person. The 'average' is lowest in the 3rd and 4th age group where it is about one per person. Thus in the Dvorine at the points where the greatest number of misreadings are found they are on an average two misreadings higher than were found in comparable groups in the Ishihara.

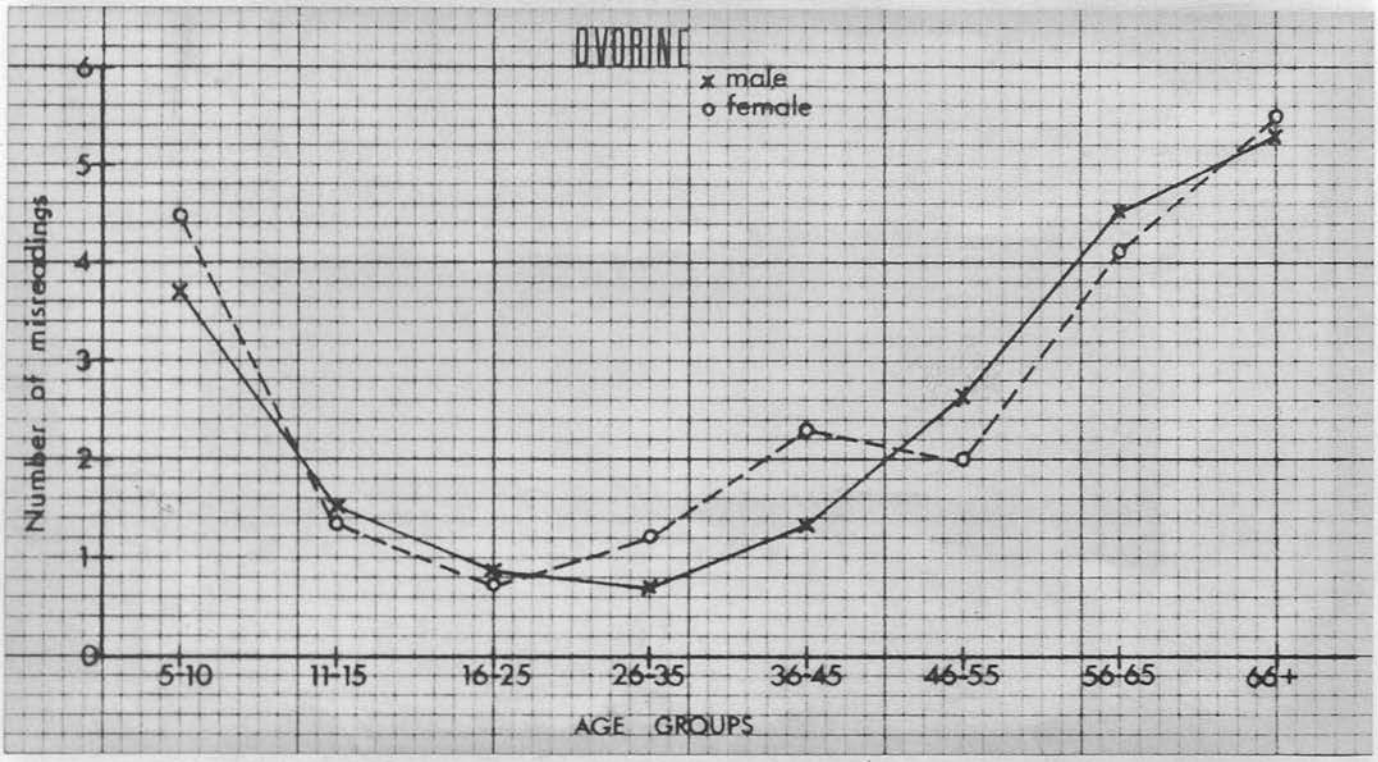


Fig. 16

If the total percentages of misreadings per age group are considered, it is seen that the very young make altogether 30% of the misreadings and the 7th age group makes 30% misreadings, whereas in the 15 - 45 age groups the mean percentage is approximately 8 to 9%.

If we compare this with the Ishihara, the percentages at the extremes of this distribution are two or three times higher in the Dvorine.

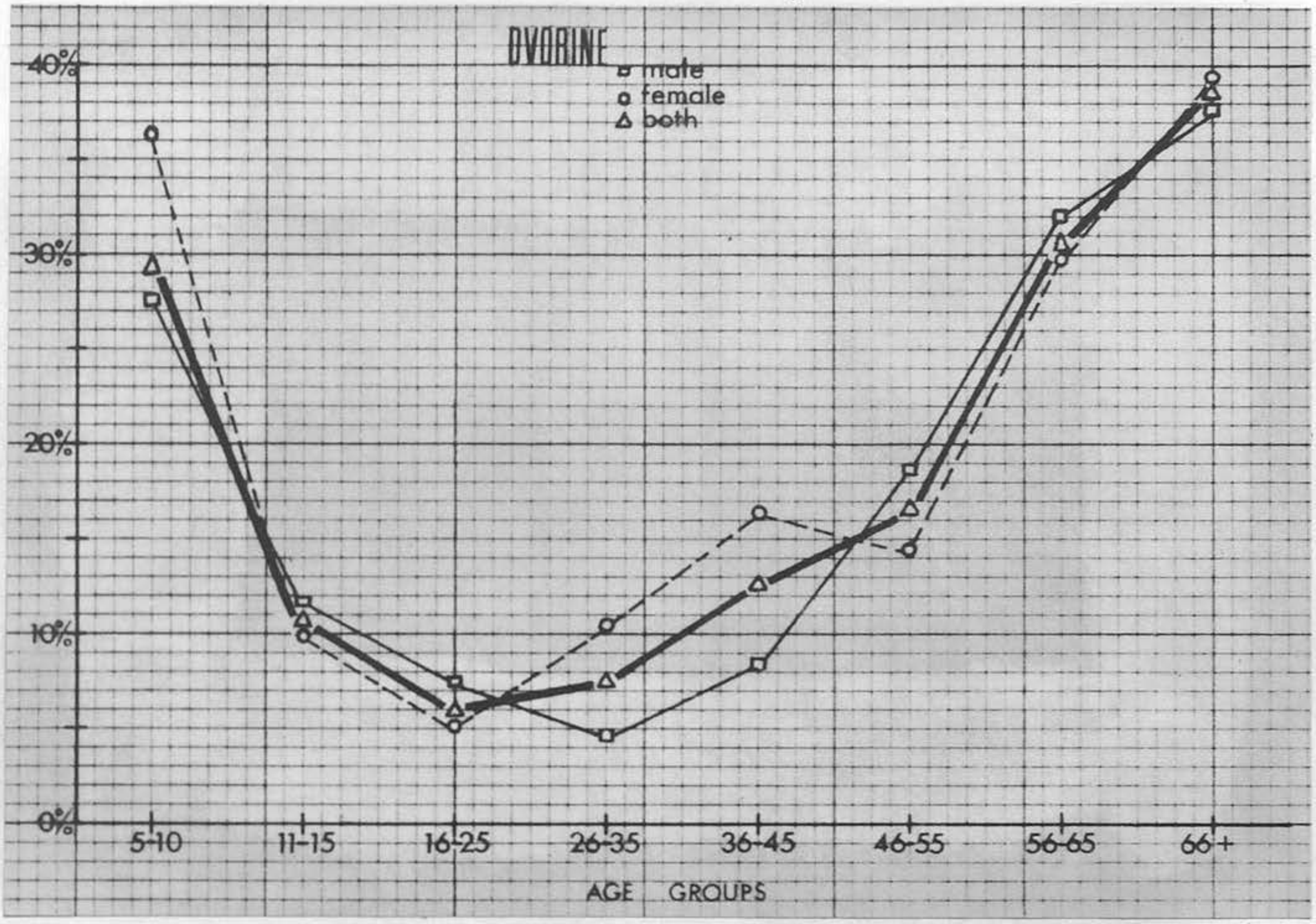


Fig. 17.

There is one other difference, the two curves here are essentially U curves, whereas the curves of the Ishihara results showed a gradual downwards slope from the youngest to the 6th age group and only in the 7th and 8th age groups was there any build-up of errors. In the Dvorine we reach the lowest point of this U curve between the 16 to 35th age groups.

There are interesting differences between the number of mistakes made by the two sexes, on the Dvorine and Ishihara tests. The incidence of misreading per person on the Ishihara was <u>constantly less</u> for females than for males in all the age groups. This is not so in the Dvorine, where

DVORINE PLATES

Total frequency of misreading per plate (total population)

Plate Series	Plate Number	Correct Reading	Both Sexes Frequency	Both Sexes %	Female Frequency	Male Frequency
I	2	67	73	14.6	44	29
	3	38	123	24.6	65	58
II	4	92	137	27.4	79	58
	5	70	81	16.2	44	37
III	6	95	30	6.0	18	12
	7	26	46	9.2	23	23
IV	8	2	76	15.2	43	33
	9	74	131	26.2	76	55
V	10	62	198	39.6	107	91
	11	4	105	21.0	67	38
VI	12	28	66	13.2	36	30
	13	46	46	9.2	22	24
VII	14	7	32	6.4	14	18
	15	39	83	16.6	48	35
		Total	1227	Average 17.5%	686 55.9%	541 44.1%

Males 244)
Females 256) Total 500

Possible No. of misreadings = 7000

Average No. (Males - 2.22)
(Females - 2.68)
per person.

the average, or total percentage of misreadings is slightly less for females in the first and last two age groups whereas in the age groups in between 26 and 55 the females have a higher average of mistakes than males. If the frequencies of misreadings of males and females for the total population are compared, we find that in contrast with the results of the Ishihara, females have a higher incidence of misreadings than males. A table showing the total frequency of misreadings per plate for the total population is given on the opposite page. Plates are given their serial number, and grouped in blocks of two. The correct number for the normal observer, is quoted in the next column, followed by a column giving the frequency of misreadings for each plate. Then there is a column giving the percentages that these numbers represent out of the total number of misreadings of all plates taken together. Lastly, the two columns on the far right give the number of misreadings for females and males for each individual plate.

The total frequency for females is 686 and for males 541, which represents percentage misreadings of 55.9% and 44.1% respectively for females and males. The 11.8% difference is significant at the 0.01 level, and reverses the trend found in the Ishihara. It can be seen from this table that these discrepancies between the sexes apply to series II, IV and V. As was found in the Ishihara these series intended for detecting red-green dochromats do not show significant sex differences in the total number of misreadings (note that major defectives were excluded). We may recall that series II requires discrimination between brown and red- series IV between orange and yellow - and series V between yellow and yellow green.

Using the X^2 technique to test the significance of the difference between males and females for the three series we obtain the following results :-

Series	Frequency of misreadings Female	Male	X^2	P
II	123	92	4.46	0.05
IV	119	88	4.6	0.05
V	174	129	6.66	0.01
Total Test	686	541	8.5	0.01

Table No. 5 Chi-square for the frequency of misreadings by males and females.

Differences in all three series are either significant at the 5% or 1% level. Unlike the Ishihara, however, (where, in the last analysis, differences between the sexes was only significant in series V) in the Dvorine, not only are the differences significant in 6 out of 14 plates, but there is also a consistently large number of misreadings by females for each of the other plates, though this may not be significant at the statistical levels usually quoted.

Finally, here is a set of diagrams of the frequencies of percentage misreadings per individual plate.

In series I (both double digit plates), the second plate is misread almost twice as often as the first, and is a more sensitive measure of age variations, for though both plates are misread by almost 30 to 40 percent of those in the older age groups, plate No. 3 is also misread quite frequently by the 5 to 10 year olds. The most frequent misreading in the first plate of this series is 57 for 67, the incidence of such misreadings varying slightly from age group to age group.

Age group	1	2	3	4	5	6	7	8
Read as 57	4.5	6.9	6.7	1.4	8.6	8.3	8.7	11.5

Table (in percentages) of the incidence of misreading for plate No. 3.

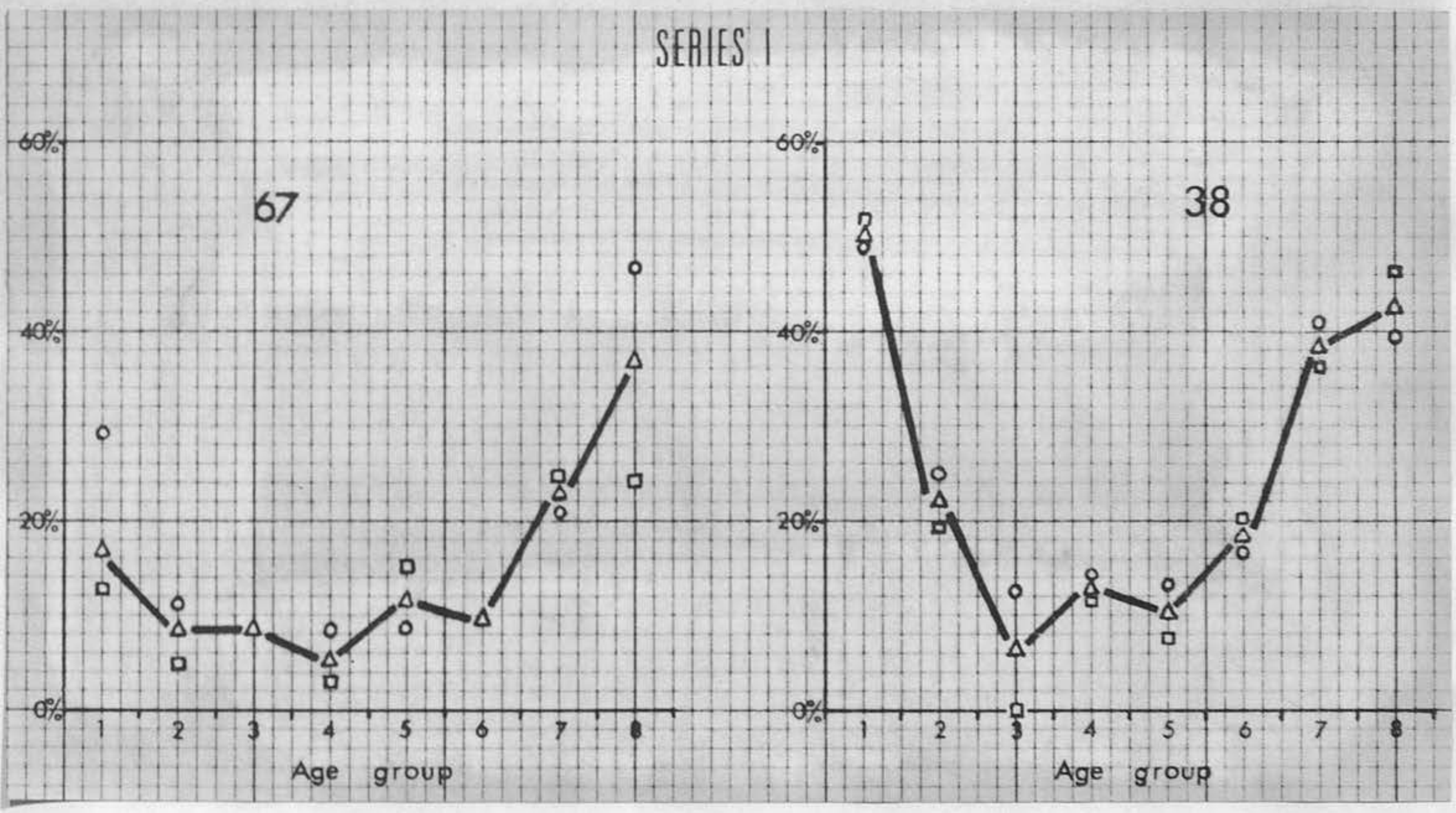


Fig. 17_2 - Frequency of misreadings per individual plate - series 1.

In series II, the first plate is misread most frequently - over $1\frac{1}{2}$ times as often as the second plate. About 36% of all misreadings for plate 4 are where 92 is read as 42. There is no age pattern in this misreading since it was misread equally often by each age group, - that is between 7 and 9 times.

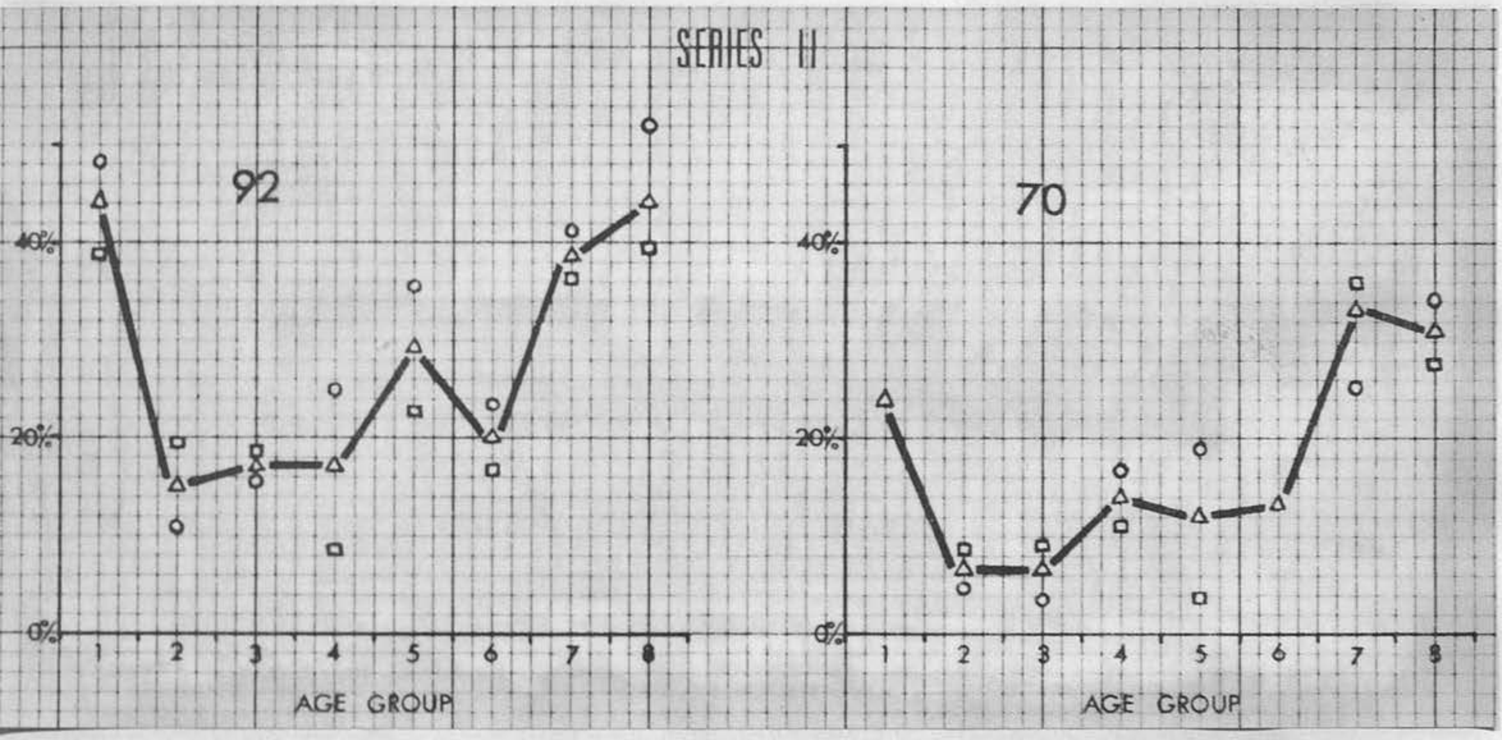


Fig. 18 - Frequency of misreadings per individual plate - series 2.

Series III is the diagnostic series differentiating between the types of dichromats. The total number of misreadings here is the lowest of all the plates in the Dvorine, especially for the plate reading 95, which was only misread 30 times in testing the 500 subjects. The age 'effect' upon the reading of these plates can be seen, though it is slight and only apparent in the second plate of this series.

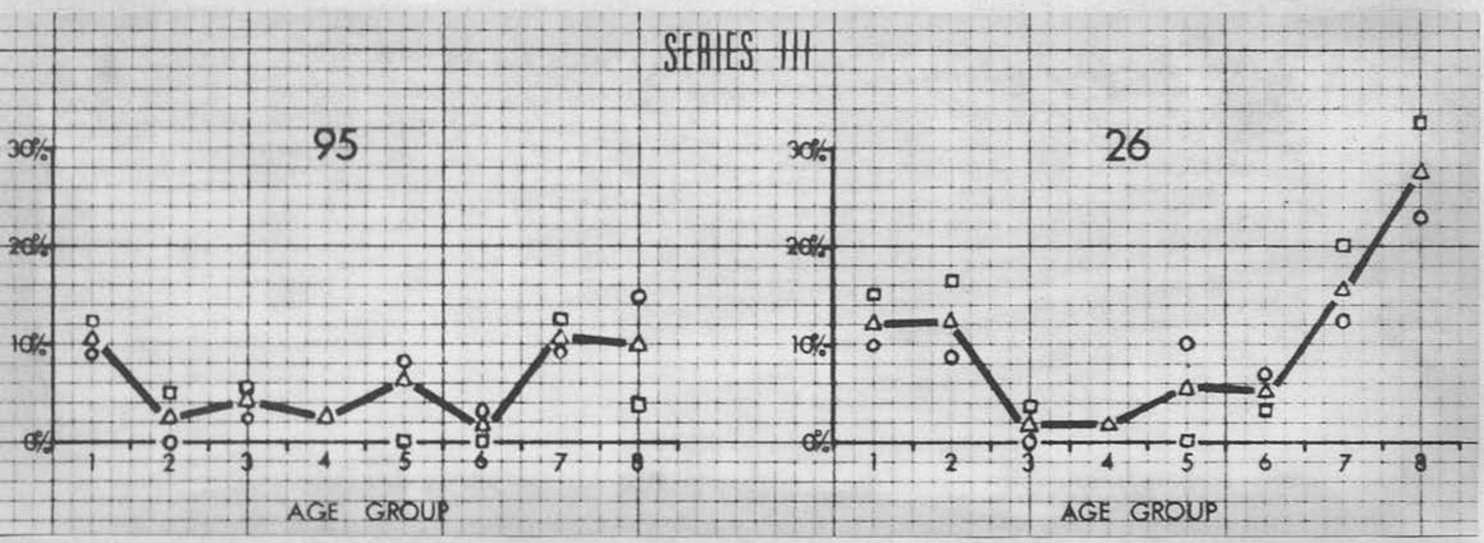


Fig. 19 - Frequency of misreadings per individual plate - series 3.

Series IV has one of the more sensitive plates for testing the influence of age on the reading of pseudo-isochromatic plates. This is the second plate of this series and produces a V type of frequency curve for the various age groups.

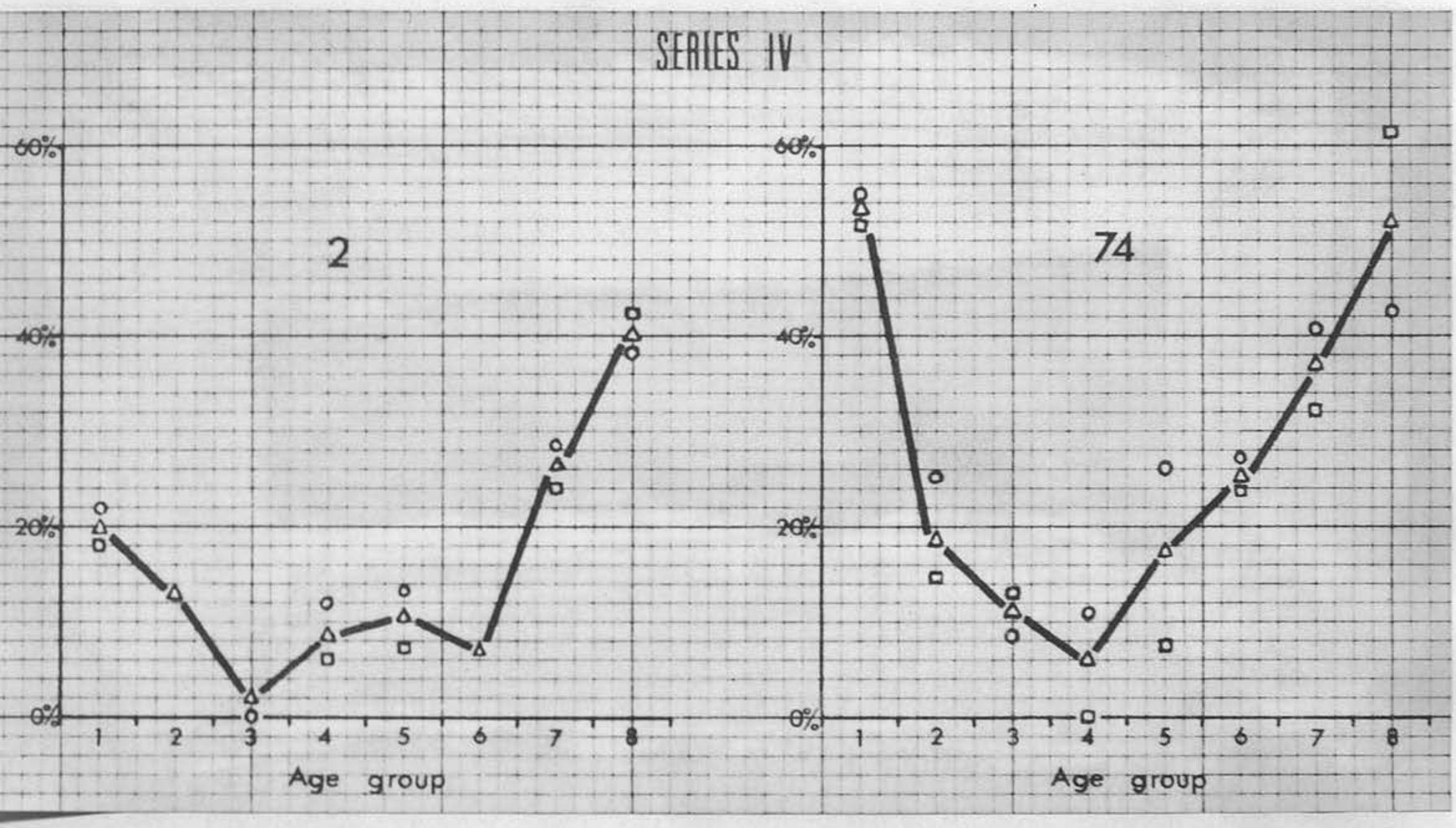


Fig. 20 - Frequency of misreadings per individual plate - series 4.

In the next series (V) the plate reading 62 for the normal observer shows the highest incidence of misreadings for the whole test and produces the sharpest 'V' type of curve of all the plates. From the position showing the lowest frequency of misreading (for age group 3) the incidence rises steeply both in the younger and in the older age groups, reaching a point where among the 65 + subjects only about 20 read it correctly.

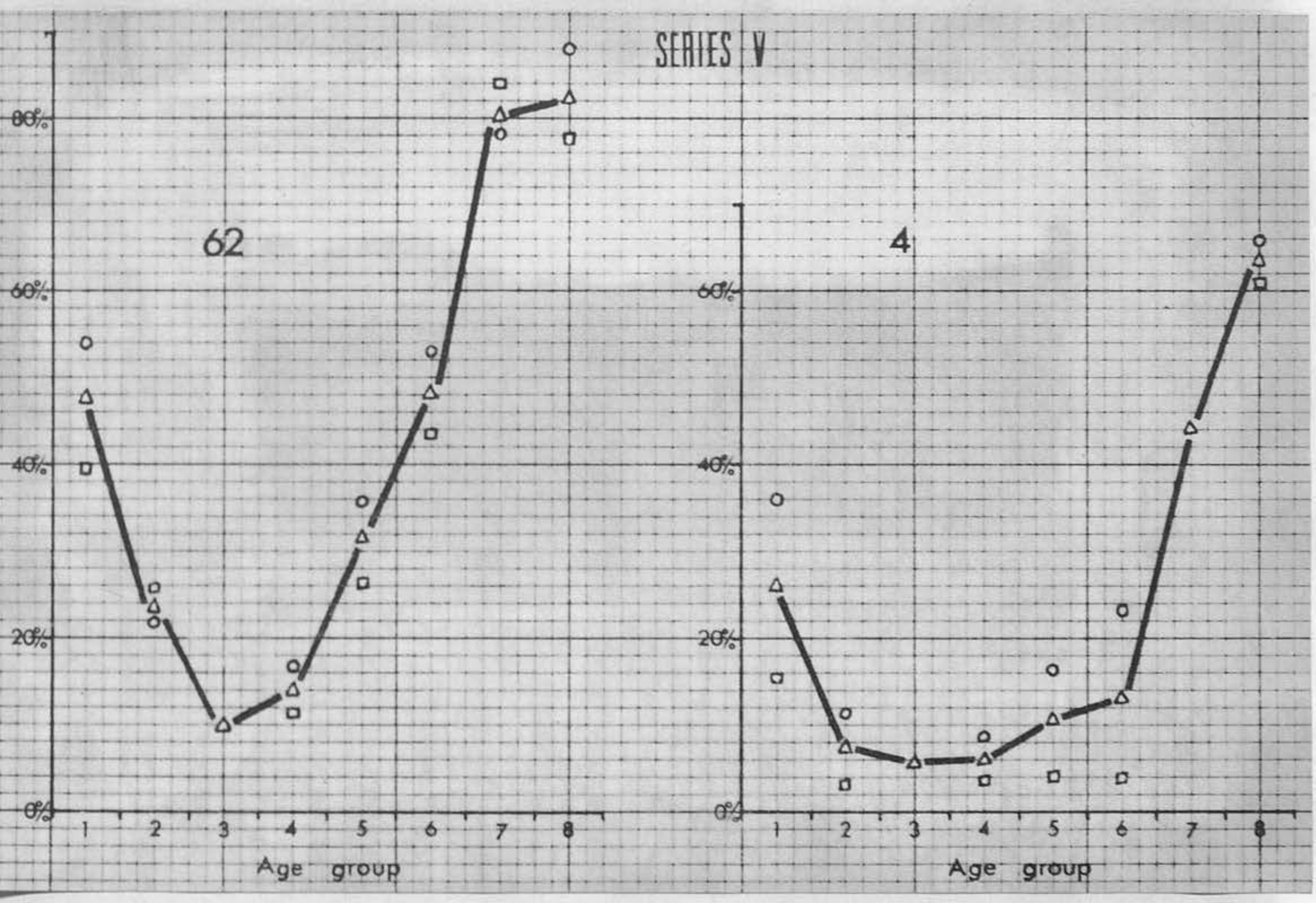


Fig. 21 - Frequency of misreadings per plate - series 5.

The last two series (VI and VII) are less affected by the 'age' variable, though a 'U' type of frequency distribution still occurs.

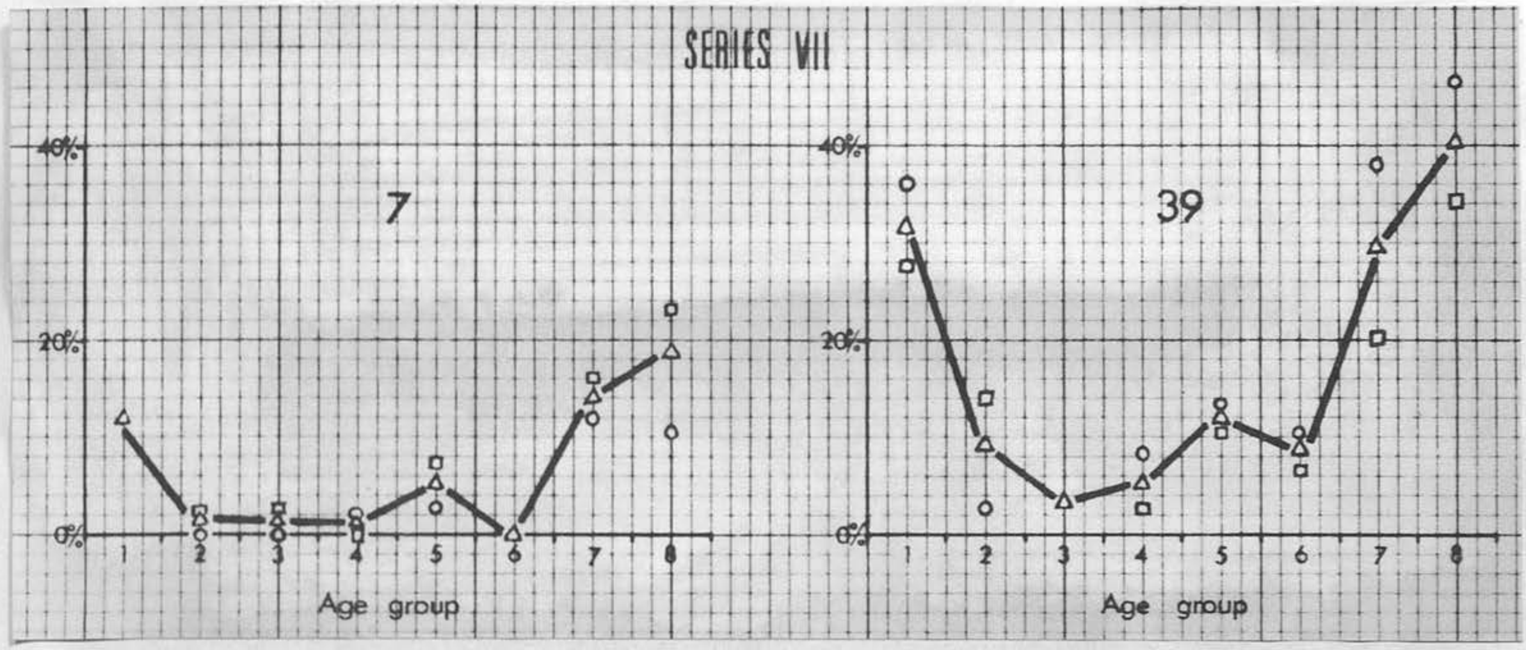


Fig. 23 - Frequency of misreadings per plate - series 7.

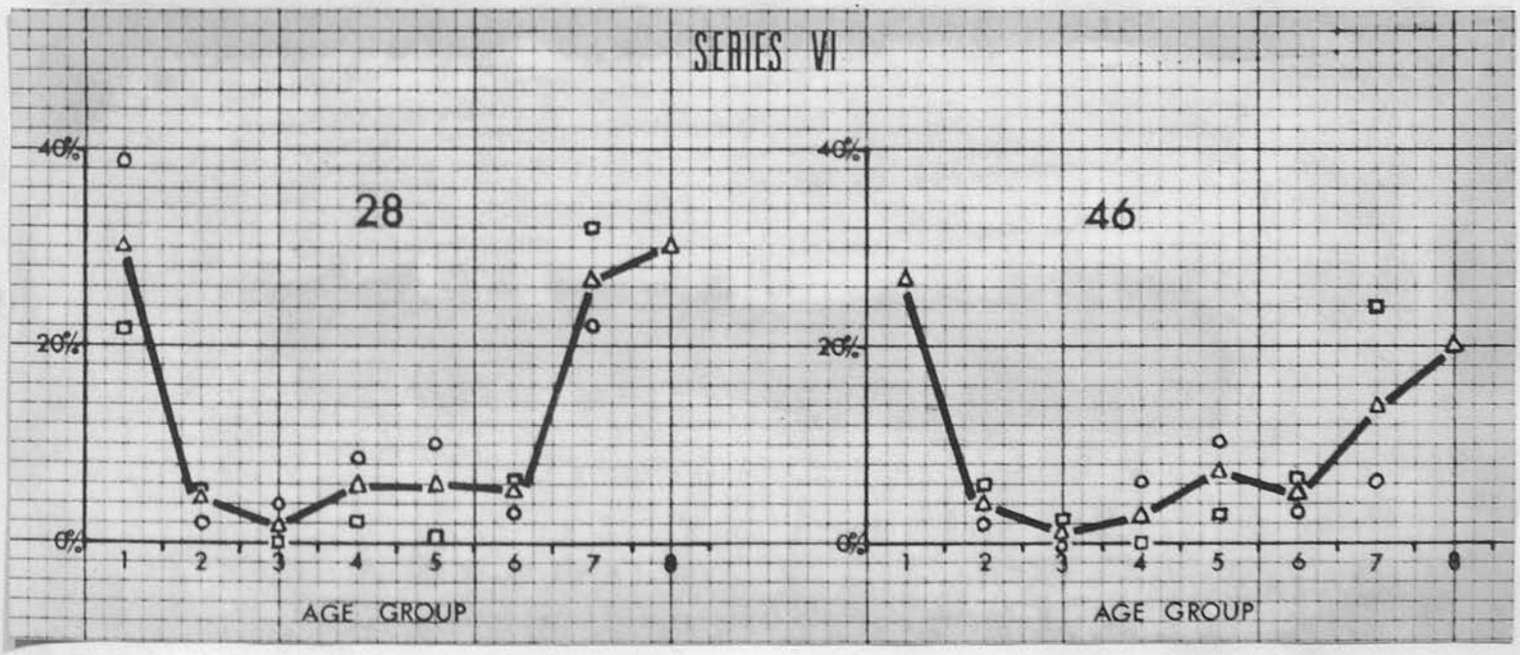


Fig. 22 - Frequency of misreadings per plate - series 6.

Total number of non-readings for age population

Age Group	Reading No.	67 2	38 3	92 4	70 5	95 6	26 7	2 8	74 9	67 10	4 11	38 12	46 13	7 14	39 15	Total per Age Group
5 - 10	66	6	13	17	9	3	2	10	15	19	16	6	8	7	8	139
11 - 15	72		1					3	3	6	5					18
16 - 25	64			2				1	5	1	2		1	1	1	14
26 - 35	71	1	1	2	2			2	2	1	1			1	1	14
36 - 45	58	1	1	5	3	1		4	4	6	5	1	2	2	2	38
46 - 55	60	1	1	2	5			6	4	10	7				1	37
56 - 65	57	6	11	6	6	3	3	14	14	27	20	8	5	3	7	133
66 +	52	9	14	8	7	3	3	16	16	31	31	9	6	6	8	167
Total per plate		26	42	42	35	10	8	56	63	101	87	24	22	20	28	

As was found with the Ishihara, some of the plates in this test are hardly affected by age variations, while others are extremely sensitive to them. However, there are far more of these plates in the Dvorine than in the Ishihara. Those least affected are the two 'qualitative' plates for the diagnosis of the deutan and protan types of defect. Plates Nos. 9, 10 and 11, especially, show a high degree of variation due to age. Most of the frequency curves are either of the U or V type, showing the lowest incidence of misreadings in either the 3rd or 4th age group (that is between 16 and 35 years).

Where there is a distinction between one digit and two digits plates in the same series (for example in IV, V, VI, VII) the plates with the double digit figure are misread more often than the single digit figures, and the double digit figures are more affected by age variation. As this test has neither the vanishing digit type of plate nor any transformation series it is difficult to make a comparison between the misreadings made by the major defectives and those made by persons in the various age groups. On the whole people who have red-green anomalies do not read the plates at all, except for the two qualitative plates. Whereas with the Ishihara it was a rare occurrence not to read plates, it is quite a frequent occurrence on the Dvorine and this again shows an age variation. The table on the opposite side gives a scatter of the percentages of non-reading of a given plate per age group. It will be seen that the middle plates i.e. plates 8, 9, 10 and 11 are the ones which have the highest overall incidence of non-reading and also the greatest variation due to age. The plates least affected are Nos. 6 and 7 for distinguishing between protanopia and deuteranopia.

Most of the other misreadings for the various plates are of the type where the first digit is read and the second is misread (or the reverse). Table No. 6 below summarises the most frequent type of misreading for the 14 plates and gives the percentage this represents of the total number of misreadings for the whole test (1227)

Plate		Misreading			Non-reading	
Serial No.	Reading	Read as	Frequency	% total per plate	Frequency	% total of mis-readings per plate
2	67	57	33	45.2	26	35.6
3	38	36	22	17.8	42	34.1
4	92	42	50	36.4	42	30.6
5	70	76	19	23.4	35	43.2
6	95	75	3	10.0	10	33.3
7	26	25	12	26.0	8	17.3
8	2	8	9	11.8	56	73.6
9	74	24	17	12.9	63	48.0
10	62	52	25	12.6	101	51.0
11	4	-	-	-	87	82.8
12	28	26	13	19.6	24	36.8
13	46	45	8	17.3	22	47.8
14	7	-	-	-	20	62.5
15	39	29	8	10.0	28	33.7

Table No. 6

The comparison of the data on the Dvorine is concluded with a comparison of the frequencies of those making 1, 2, 3, 4, etc. mistakes for the various age groups. Again the 'age' population is divided into four sub-groups, each containing approximately 120 subjects.

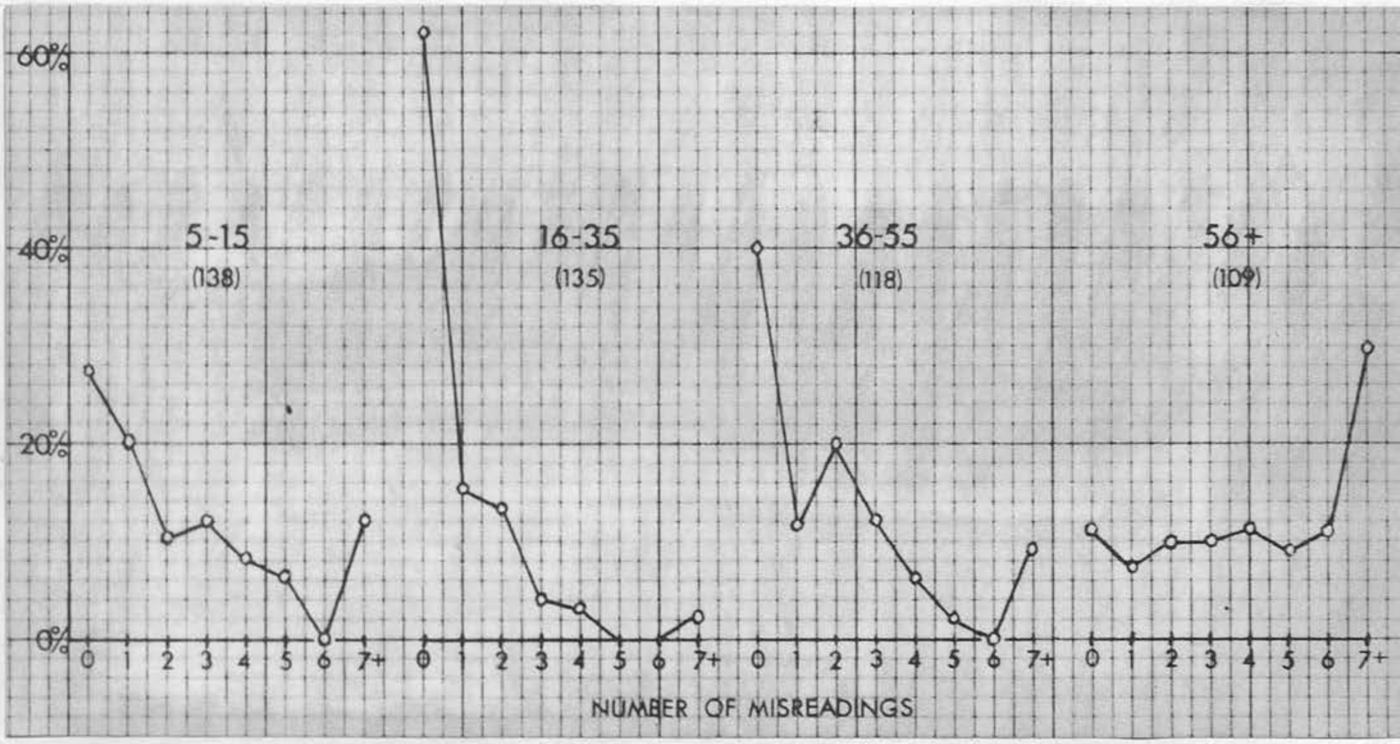


Fig. 24 - Frequency of misreadings per age-group.

The diagrams above show the frequency distributions for these four sub-groups, and is constructed along lines similar to the comparable diagram used in the Ishihara. As before, all major red-green defectives have been excluded. Only in one age group in the Dvorine (that is the 16 - 35 group) does the frequency curve resemble a reversed J type of curve. The other age groups produce a fairly irregular curve, and finally in the oldest sub-group the picture is almost the reverse of the curve shown by the 16 - 35 group. Here the greatest frequency is that of people making 7 or more mistakes, (reaching as much as 30 per cent) whereas those who read the test faultlessly contribute only 11 per cent of this group. In this respect the curve for the last age group on the

Dvorine is quite unlike the comparable one for the Ishihara test where there were still larger numbers of people with no mistakes than there were with more than one misreading.

Results of the Dvorine test for the population of this research will now be compared with the data of other homogeneous groups. In addition to the previously mentioned data compiled by Belcher, Sloan, and Crawford, the results of G. E. Peters' study (1955) on 800 male police candidates will be added. Attention is drawn to the fact that this is the only 'closed' population to be used in the present comparison of performances on the Dvorine. The age span in this group was from 20 - 42 with a mean age of 29 and an S. D. of 4 years, - thus two thirds of this population were between the ages of 25 - 33, that is they are roughly comparable to our 16 - 35 age group.

Researcher	Year	'Inclusive' Male	'Inclusive' Female	'Inclusive' Total	'Modified' Male	'Modified' Female	'Modified' Total	Type	Degree of Randomness
Belcher etal	1958	468	32	500			462	Students & Staff	Open, random
Crawford	1955				24	41	65	1st Ord. Psych.	Selected
Lakowski	1958				244	256	560	Age Groups	Open, random
Lakowski	1962			396	-	-	364	1st Y. Psych.	Open, random
Peters	1956	800		800	745	-	745	Police Force	Closed, random
Sloan	1955						100	Not known	Selected

It is doubtful if the student population of this research can be looked upon as a 'closed' type of sample. Data was collected over a period of 3 years - where in the years 1960, 1961 only 80 or so subjects were tested on the Dvorine and in 1962 over 200 were tested to bring the number of total tested on the Dvorine to round about 350. Although in 1960/61 only small numbers were tested on the Dvorine the testing was random in that each student was tested on whatever test was available (i.e. either Ishihara or Dvorine). However, in some respects this population is more a 'closed one in the sense described in the section dealing with the Ishihara test, than the population of Sloan or Belcher. The above considerations are very important, because they might explain why the populations of Belcher, Sloan and Crawford produce types of frequency distribution so different from the 1962 student populations in this research, and the population tested by Peters.

A table hereunder gives the number of frequencies of misreadings for the 1962 student population.

Student Population - 1st Ord. Psych. 1959-1962 N 397

Number of mistakes	Total Subjects Number	%
0	162	40.9
1	89	22.4
2	60	15.0
3	27	6.8
4	10	2.5
5	7	1.7
6	5	1.2
7	4	1.0
8	2	0.7
9	5	1.2
10	-	-
11	2	0.7
12	1	0.3
13	4	1.0
14	18	4.5
Total	396	99.9%

The diagrams showing the frequencies of number of mistakes for the three 'inclusive' populations shows a close correspondence in the frequency of those making 7 or more mistakes on this test. There is, however, a great discrepancy between the frequencies of those with perfect readings in these populations :-

Belcher 83.6%, Lakowski 40.9%, Peters 56%

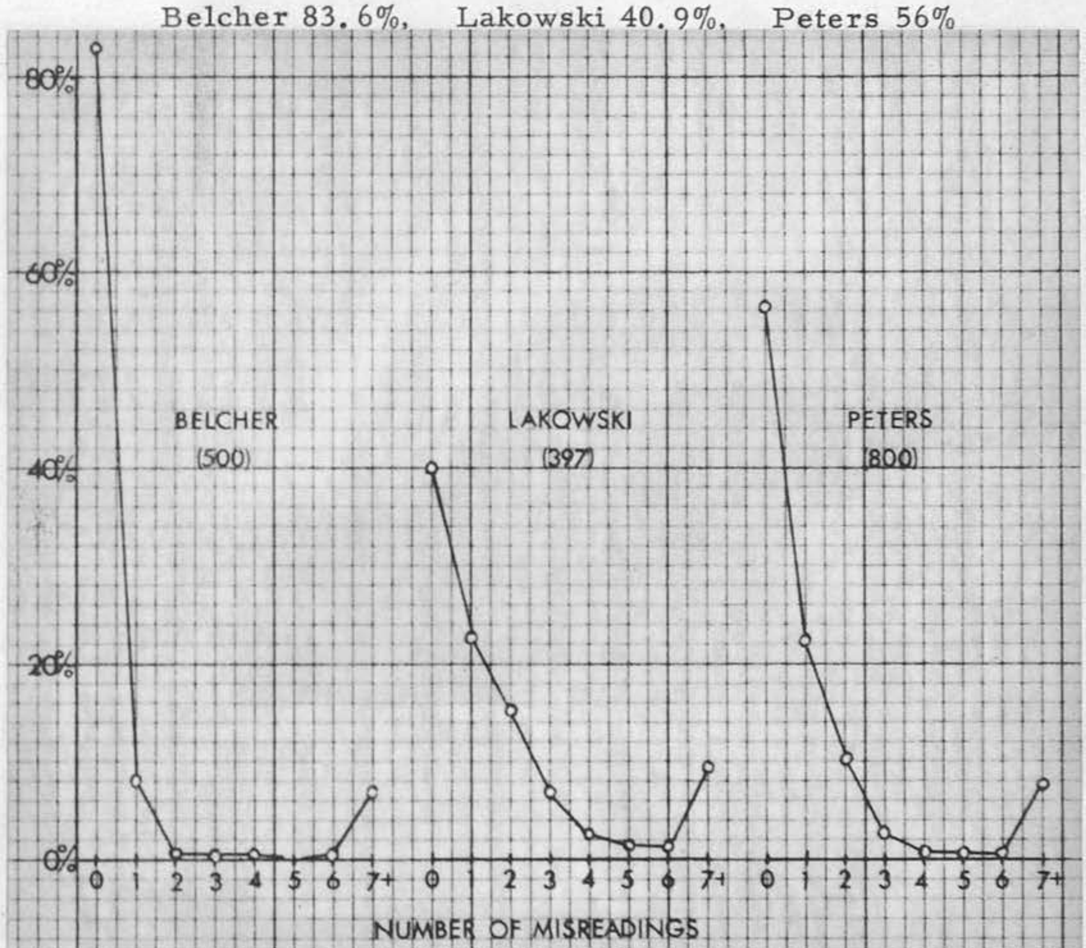


Fig. 25 - Frequency of misreadings - three student populations.

The first percentage is significantly higher from those of Lakowski and Peters. If we look at the age sub-group curves for the 16 - 35's it will be seen that the percentage for those with no errors is 62% (which is very close to the figures quoted by Peters).

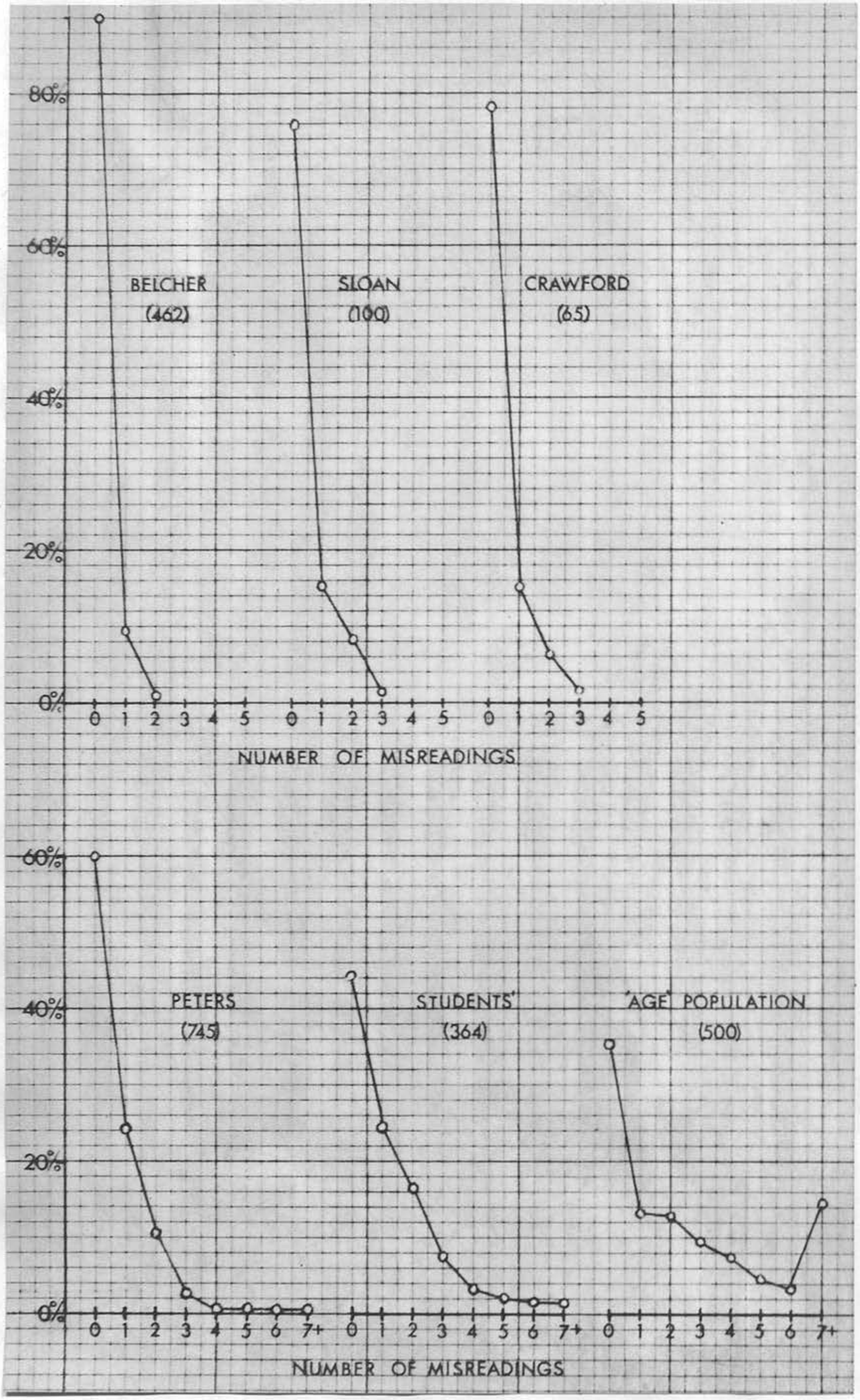


Fig. 26 - Frequency of misreadings - six 'modified' populations.

The diagram giving the six 'modified' populations consists of samples from which major red-green defectives were excluded. Notice that the frequency curves are again very similar to the populations of Belcher, Sloan and Crawford in both the steep decline and suddenness with which these curves terminate at the 3 mistakes position. The curves for the three other populations i.e. (Peters', Students' and the total age population) show a gradual sloping downwards and a lower percentage of those with no mistakes, and those with 1, 2, 3 mistakes, than were seen in the three previous curves. In another respect, there is a difference - even in Peters' population only a small incidence of people make 4, 5 or 6 and 7 mistakes. In this respect the Age Population curve is different from all the other five, having a much higher percentage of people making 3, 4 or 5 mistakes, and in addition showing an incidence of 14.5% making 7 or more mistakes (red-green defectives excluded).

Some research workers have suggested that certain types of digits are read more easily than others. The Dvorine test should allow us to examine this hypothesis as the digits of the figure and ground are made mostly of dots of the same hue and vary only in saturation. It is possible therefore to find out whether some of the 'digits' are more difficult to read than others. Only plates that have a high percentage of misreadings (above 20%) were chosen. Table No. 7 gives digit analysis for 4 plates, showing the effect of age upon this type of misreading.

	Series I				Series II		Series V	
	Plate 2		Plate 3		Plate 4		Plate 10	
Age Group	6	7	3	8	9	2	6	2
1	4	1	0	13	7	3	7	1
2	6	0	0	9	8	2	7	5
3	3	0	2	2	7	1	5	0
4	2	0	1	5	9	1	6	1
5	6	0	0	3	11	0	9	3
6	5	0	2	5	9	0	13	2
7	6	1	2	6	11	0	11	2
8	8	-	3	6	12	1	9	0
Total	42	4	10	49	73	8	67	14
Total Misreadings per plate	73		123		123		198	

Table No. 7

The table below (No. 8) gives the frequency of misreadings for a further three plates (showing only slight effect due to age).

Series II Plate 5		Series IV Plate 9		Series VI Plate 12		Plate
7	0	7	4	2	8	
13	31	28	31	5	40	Digit misreadings
137		131		66		Total misreadings

Table No. 8

To obtain an idea of the level of significance for these types of misreadings consult table No. 9

Using these tables, we can examine our initial hypothesis, by finding plates where one of the numbers of the composite figure occurs in more than two plates. Firstly, however, let us see the type of information that can be obtained from these tables.

Take plate No. 2 of the I series, reading 67. At the bottom of the column for this plate the number of total misreadings is shown - in this case it is 73. In the previous line (headed "Total") the number 42 refers to the total number of misreadings for all age groups of the first digit (6) of plate two, when the second digit was read correctly. In a similar manner the total 4 represents the number of misreadings of the second digit (7) for all the age groups where the first digit was read correctly.

In plate No. 3 - (correct reading 38) the total incidence of misreadings for the whole population is 123. The digit '3' is misread alone 22 times, while

Frequency of Misreading

Series	Plate No.	Reading	First Digit	Second Digit	X^2	P
I	2	67	42	4	29.76	0.001
	3	38	10	49	25.78	0.001
II	4	92	73	8	50.50	0.001
	5	70	13	31	7.4	0.01
IV	9	74	28	31	34.68	0.001
V	10	62	67	14	25.69	0.001
VI	12	28	5	40	25.67	0.0001

Table No. 9

Table No. 9 gives chi-square scores for the frequency of misreadings of the first in comparison with the second digit of a given plate (Yates correction factor of 0.5 was used where appropriate), 1 d f.

the digit 8 is misread alone 48 times. Table No. 7 gives the frequency misreadings of the digits for the eight age groups while the next table refers only to digit misreadings of the total population. In these plates (i.e. plates No. 5, No. 9, No. 12) significant differences in the type of misreading from age-group to age-group appear.

Comparing all the plates in the two tables we notice that some numbers have a higher frequency of misreading than others - the numerals 9, 6, 8, 0, 4 have the highest frequency of misreadings, while the numbers 2, 3, 7 have the lowest.

If these frequencies are expressed as ratios of each digit misreading to the total number of misreading of digit 1 and 2 for that plate :-

$$\frac{\text{1st Digit}}{\text{1st + 2nd Digit}} \quad \text{and} \quad \frac{\text{2nd Digit}}{\text{1st + 2nd Digit}}$$

the following ratios are obtained in order from the highest to the lowest ratios :-

Numeral	9	6	8	0	4	7	3	2
Ratio	1.1	1.15	1.15	1.4	1.9	5.52	5.9	7.9

Four of the numerals (Nos. 2, 6, 7 and 8) appear in more than one combination.

Numerals	No. of Plate	Reading	Ratios
8	12, 3	28 & 38	1.1 & 1.2
6	2, 10	67 & 62	1.1 & 1.2
7	2, 5, 9	67, 70 & 74	11.5, 3.3 & 2.1
2	4, 10, 12	92, 62 & 28	8.1, 5.7 & 9.0

The numerals 8 and 6 have low ratios on two separate plates while 7 has low ratio in combination with the numeral 4, but a high ratio in combination with numerals 6 and 7. Numeral 2 consistently has a higher ratio than the second digits, it occurs with, - 9, 6 or 8.

Thus on the Dvorine the numeral '2' is misread less often than any other numeral it occurs with in three plates, while the numer '6' is misread more frequently in each situation and also has the highest percentage of all misreadings for any of the plates mentioned here. It may safely be concluded that as they are used in the Dvorine, certain numbers are easier to recognisethan others. However, one has to be very careful in making such inferences from the results quoted here since it is possible that in the plates with a high frequency of misreadings the distribution of the particular coloured dots of which the numeral is composed is such that at certain vital points, there is a greater affinity between the colour of the figure and the neighbouring coloured dots of the background.

II. Summary and Conclusions - The Dvorine test shows variations in the number of misreadings found among subjects of different ages - the very young and the very old have a higher incidence. The resulting curves for these frequencies are of four types. They are either 'V' shape, 'U' shaped, or show an increased incidence in the older age groups or, lastly, some plates show no variations at all throughout the age groups.

More plates show variations from age to age than were found in the Ishihara. Conversely, fewer plates are unaffected by the age variable and these are mainly the plates for detecting dichromats. The frequencies at the end-points of the age distributions are higher for the Dvorine both in terms of average per person per age group, and in terms of total percentages of misreadings. On an average the frequency of mistakes made was two fewer in the same age groups in the Ishihara than it is in the Dvorine.

The non-congenitally determined colour blind do not make misreadings that are specific to such groups. On the Dvorine the most frequent type of error is non-recognition of the numbers. This type of error is particularly frequent in the first and in the last three age groups.

The distribution curve of those making 1, 2, 3 and more mistakes shows a much higher frequency in the 5 - 15 age groups in the Dvorine than it did in the Ishihara. The frequency curves for both tests are similar for those between 16 and 55 but differ remarkably for those in the 56 plus age distribution. In this group the smallest group in the Dvorine is of those making no mistakes while the incidence of those making 1, 2, etc., mistakes increases, till it is highest for those making

7 + mistakes.

Generally speaking it can be said that the Dvorine test is a more sensitive measure of changes in colour discrimination in the various age groups - than the Ishihara test. It is no exaggeration to say that if plates Nos. 8, 9, 10 and 11 in the Dvorine are misread the subject must either be at least 45 years of age, or have some form of acquired dyschromatopsia.

(c) Tritan Plates

In addition to the pseudo-isochromatic plates, plates of the tritan type were also employed in this research. These consisted of individual plates designed by Farnsworth and Willmer. The Farnsworth plates used were F. 2, F. 3 and F. 5 and Willmer's W. 2 and W. 11 plates were also used.

I. Description - The F. 2 and F. 3 plates refer to the two plates of the same design where F. 2 was the first designed and F. 3 was an improvement on F. 2 in its colour selection. This plate (i. e. F. 3) consists of a square background of $\frac{1}{4}$ in. spots of two shades of mauve within which two squares are superimposed on each other one in shades of green and the other of two shades of blue. Two of the colour spots, however, belong to both squares. F. 5 consists of olive, light green, dark green and yellow dots which make up the figure and the background is composed of light purple, dark purple, dark blue and light blue dots. The normal person reads the figure 5 whereas the tritan subject is supposed to read either 2 or nothing.

The two charts W. 2 and W. 11 were constructed by Dr. Willmer of Cambridge. W. 2 shows a number 2 made of green dots on a blue and mauve background, which people with normal vision are supposed to be able to recognise when the chart is shown at a distance ofnot more than 15 ft. Tritanopes will not read the number 2.

The other chart, W. 11, consists of blue and green squares forming a number 11 on a background of several shades of mauve, yellow and orange squares and it is supposed that people with normal vision will not read it within about 16 ft.

but that it becomes clearly discernable at greater distances. Tritanopes are said to be able to read it at quite close range.

In the past, these plates were not often used and none were used for mass screening, therefore, little data is available. Occasionally, however, they were used for studying individual tritanopes. Kalmus (1955) used them in his study of the familial distribution of congenital tritanopia where he tested about 50 normal subjects and 16 diagnosed (i.e. colorimetrically assessed) tritanopes.

II. At this point a short history of these plates is necessary as those used here are actually reproductions cut from illustrated magazines. Farnsworth's plate F.5 was obtained from the Picture Post Weekly where it appeared in an article on Colour Vision in 1952. The other plates, except for Farnsworth's F.3 plate, that is, Willmer's W.2 and W.11 and Farnsworth's F.2 plates were obtained from an article by Kalmus in the Journal of Human Genetics. The improved Farnsworth F.3 plate was obtained from the late Commander Farnsworth, but will not be analysed here as the plate arrived in the middle of this research and data is therefore insufficient.

The vital question is, what is it that the plates are designed to test ? Do they test tritan or tetartan defects, or do they measure the amount of pigmentation in the ocular media, and in the macular region ? Let us first take Farnsworth's F.5 plate, which is read as 5 by the 'normal person'. In his paper entitled the 'Characteristics of Tritanopia' published in the Journal of the Optical Society of America in August, 1952, Wright says that in 'the Picture Post edition of May 12th, 1951 in the article by Mr. T. Morley he included a suitable test for tritanopes which was prepared some years earlier by Lt. Comm. Dean Farnsworth of the Medical Research Laboratory, United States Submarine Base, New London, Connecticut and the copy had been given to him in 1947 by Farnsworth at a colour vision conference held at Cambridge'. Then, he says, 'that the chart was designed on the same principle as the Ishihara charts but was made from a selection of Munsell colours chosen to demonstrate the colour confusions associated with the small field of tritanopia. However, the chart was also proved satisfactory in detecting congenital tritanopia'. It was this chart that brought Wright the answers in his study on the incidence of tritanopia in the population in general. He adds that in the case of the Farnsworth chart elderly observers may exhibit some of the characteristics of tritanopia because of yellowing of the eye media and the macula. Those who suggested that instead of seeing 5, they saw only 2 were looked upon as tritanopes and some of them were invited to participate in further studies.

However, in a paper by Judd, Plaza and Farnsworth entitled 'Tritanopia with abnormally heavy ocular pigmentation' the authors refer to a 'poly-chromatic' plate of the reversible figure type which was constructed and based on apparent compunctal point at 460 mu. and was effective in detecting colour deficiency in this particular subject, that they were describing. This plate was devised because the Stilling plates intended to detect tritanomaly were ambiguous and the colours used by the Rabkin tritanomalous plate ineffective in detecting this man's special colour deficiency, (as these plates were based on apparent compunctal point at 400 and 475 mu). The authors also give Munsell designations of the colours used in this plate, and these correspond with our own measurements, allowing for the fact that the plate used in this research is a copy of the original.

In private conversation and correspondence with Dean Farnsworth the writer enquired about the F. 5 plate in its original form, to which he got the following reply : (Quote from a letter dated 11th March, 1959.....), "I brought the original in 1947 to Willmer as a demonstration of the effect of macular pigmentation in the fovea. It is based on a co-punctal locus of about 460 which seems to be the dominant effect of macular pigmentation. A proper tritanomalous chart for a fairly clear eyed person would use a point of origin of about 400 mu. When Wright asked me about publication of a chart in the 'Post' I sent over a plate made especially for the purpose but, newspaper deadlines being what they are, they had to use the one that was available. (Incidentally now that you bring it up I do not know whatever happened to the real tritanomalous chart)! "

Thus the same plate is used by one author to detect ocular pigmentary changes and by the other to detect tritanomaly.

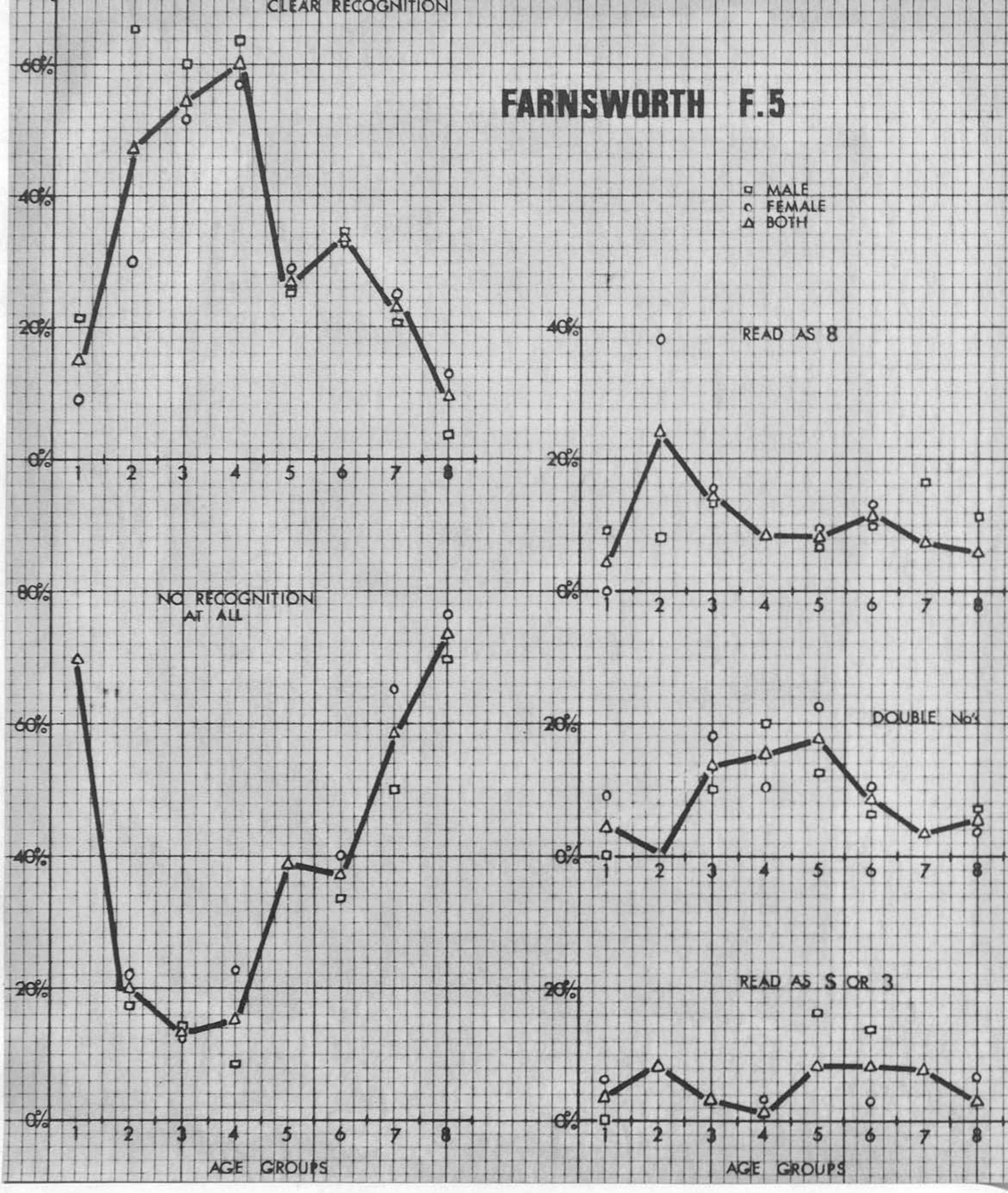
All the other plates used here are recognised by others as tritan plates. Kalmus in particular used them for detecting tritan types of defect. Dealing first with Willmer's plates, Kalmus says that the W. 2 plate is an almost perfect test for tritanopia, since 15 out of the 16 diagnosed tritanopes failed to recognise the 2 even when the plate was brought quite near to them, whereas all 52 normal subjects of whom 23 were relatives of tritanopes, recognised a clear

2 at a distance of 4 to 6 ft.

Of plate W. 11 he says that this is less effective for the purpose of detecting tritanopia. He observed that 55 normal subjects read the 11 at an average minimum distance of about 4 ft. (arranged 2 to 8 ft.) whereas for the 16 tritanopes the average minimum distance was about 2 ft. (range 1 to 5 ft.)

In the Farnsworth F. 2 chart where the subject is expected to see both green and blue squares Kalmus found that among normal subjects 9 out of 43 saw only the green square and 34 saw both green and blue squares but saw the green as the more dominant one, whereas of the 16 tritanopes 14 saw only the blue square and 2 saw both green and blue squares, but saw the blue square more distinctly. Thus none of the normal subjec ts saw the blue square alone and none saw the blue as the dominant one. The normal subjects either saw the green square alone (i. e. approximately 1/5th of the subjects) or saw the two squares but with the green square predominant.

Of the 400 subjects in this research who did this particular test, 4 subjects see the blue square alone and 9 subjects see both squares but see the blue as the dominant one. In all, 11 out of those who took this test failed it, giving a $2\frac{1}{2}\%$ incidence of subjects with a tritanopic type of reading.


FARNSWORTH F.5
CLEAR RECOGNITION
NO RECOGNITION AT ALL
READ AS 8
DOUBLE No's
READ AS S OR 3
□ MALE
○ FEMALE
△ BOTH
0%
20%
40%
60%
80%
1
2
3
4
5
6
7
8
AGE GROUPS
AGE GROUPS

III. Analysis of Results -

(i) The Farnsworth Plates are analysed first and Willmer's plates later.

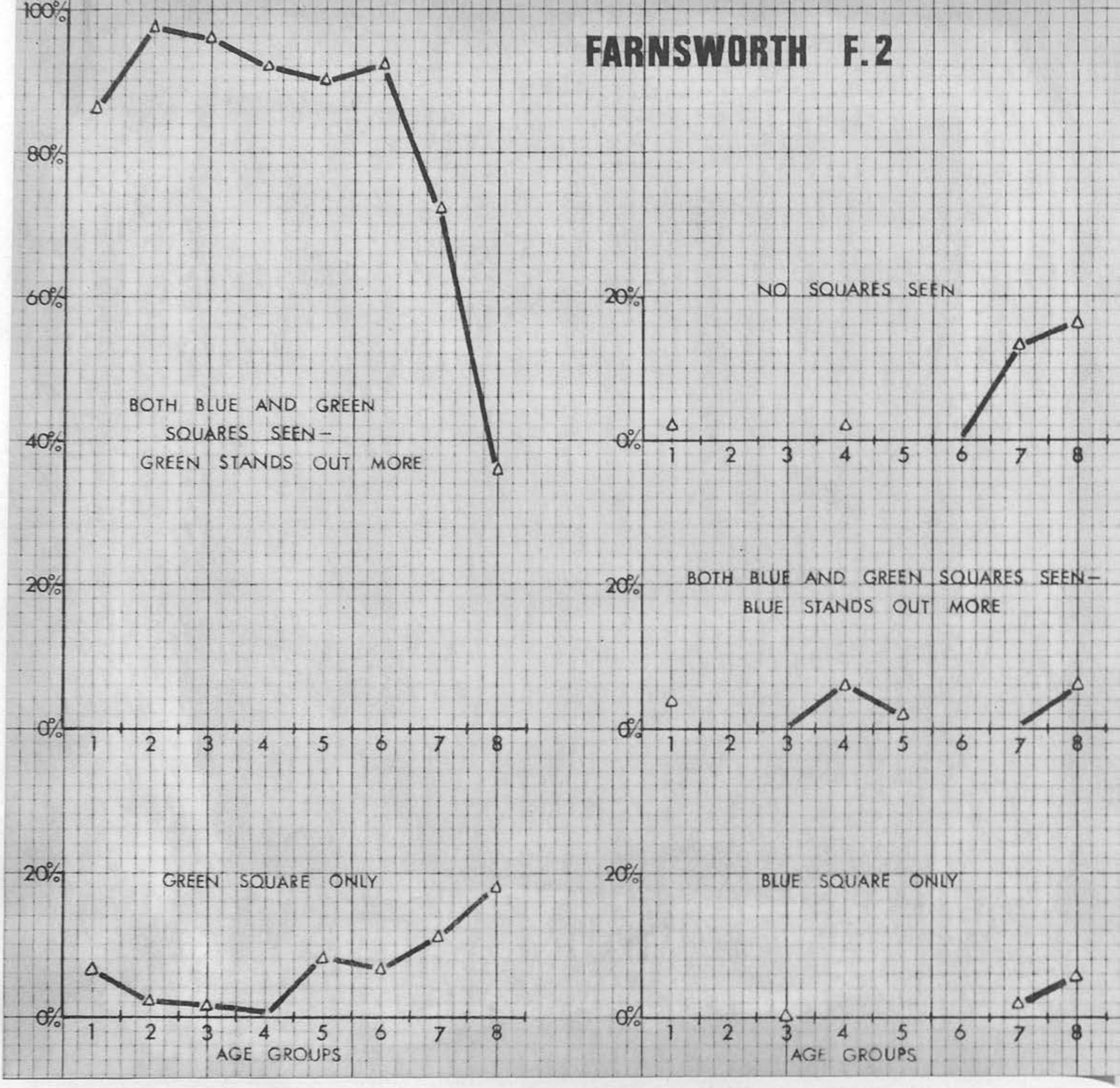
The results for Farnsworth's F. 5 plate have been classified in five sub-sections as shown in the diagramatic summary of results on the opposite page. Frequencies are given for those who saw a clear 5 without any probing, encouragement or re-reading - followed by the frequency distribution for the incidence of those who did not recognise any figure whatsoever. The third diagram shows those who read it as No. 8 and the fourth gives an account of those who read it as a double number. The final diagram quotes the incidence of those reading it as 's' or as the number three.

The curve giving the incidence of 'clear reading' shows an inverted 'v' type of graph with its peak in the fourth age group (that is 26 to 35), where approximately 60% of the people tested read the plate. From this point the curve slopes down until only about 18% in the youngest age group read it as 5 and in the last two age groups about 22 and 12% respectively, recognise it as 5.

In the diagram giving the incidence of no readings, we have a 'V' type of curve showing the highest incidence of no readings in the very young, (in fact, about 70%) and only slightly less in the 55 and over age groups. The curve is the converse of the previous curve where only about 10 to 15% between the ages of 15 and 35 failed to read it.

The other three curves show that in most age groups round about 10% read this plate as 8. The reading of double numerals is most frequent in the 16 to 40 age groups and about 5% to 6% over all age groups read the plate as 5 or 3.

FARNSWORTH F. 2



It is interesting to note that when people are tested with this plate, the reaction among the very young and also among the older subjects is one of surprise that they should be asked to look for a number when they are convinced that none exists.

When we recall that Wright used this plate for detecting tritan subjects although it was designed for the study of pigmentation in the macular region, the results become interesting. The increase in numbers of those with no readings and the decrease of those who can read it as 5 from the middle age group on, could now be explained by the progressive yellowing of the macula, but the increase, for instance of no recognition among the very young, is not so easily explained.

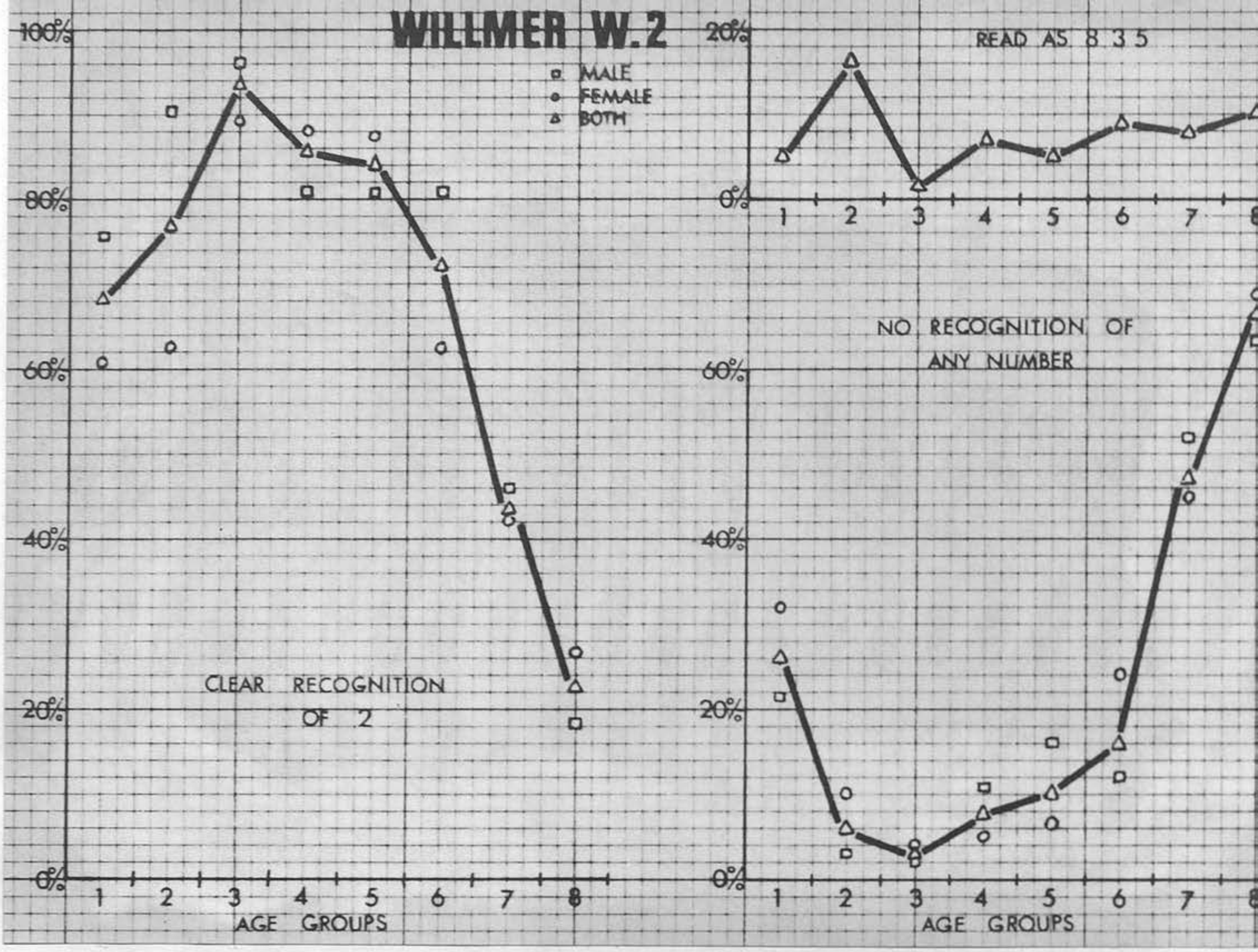
For Farnsworth's F. 2 plate we have five frequency curves.

The first shows an incidence of about 90% in all age groups (except the two oldest) who read it as a blue and green square with the green square dominant. In the 7th and 8th age group the incidence dropped to 70% and 36%.

The next diagram shows the incidence of those who saw the green square only, and here we notice that it is in the over 40 age groups that there is an increase in the incidence of these readings. Very few young people see a green square alone.

The next two frequency diagrams show that more of the older people of 55 and onwards fail to see the squares altogether (the incidence is approximately 18% to 20%).

The distribution of those seeing the blue square more clearly than the green square is irregular. Some are in the very young and middle age groups and a few in the very old group. It is also found that seeing the blue square alone occurs in about 4 subjects in each of the third, seventh and eighth age groups.

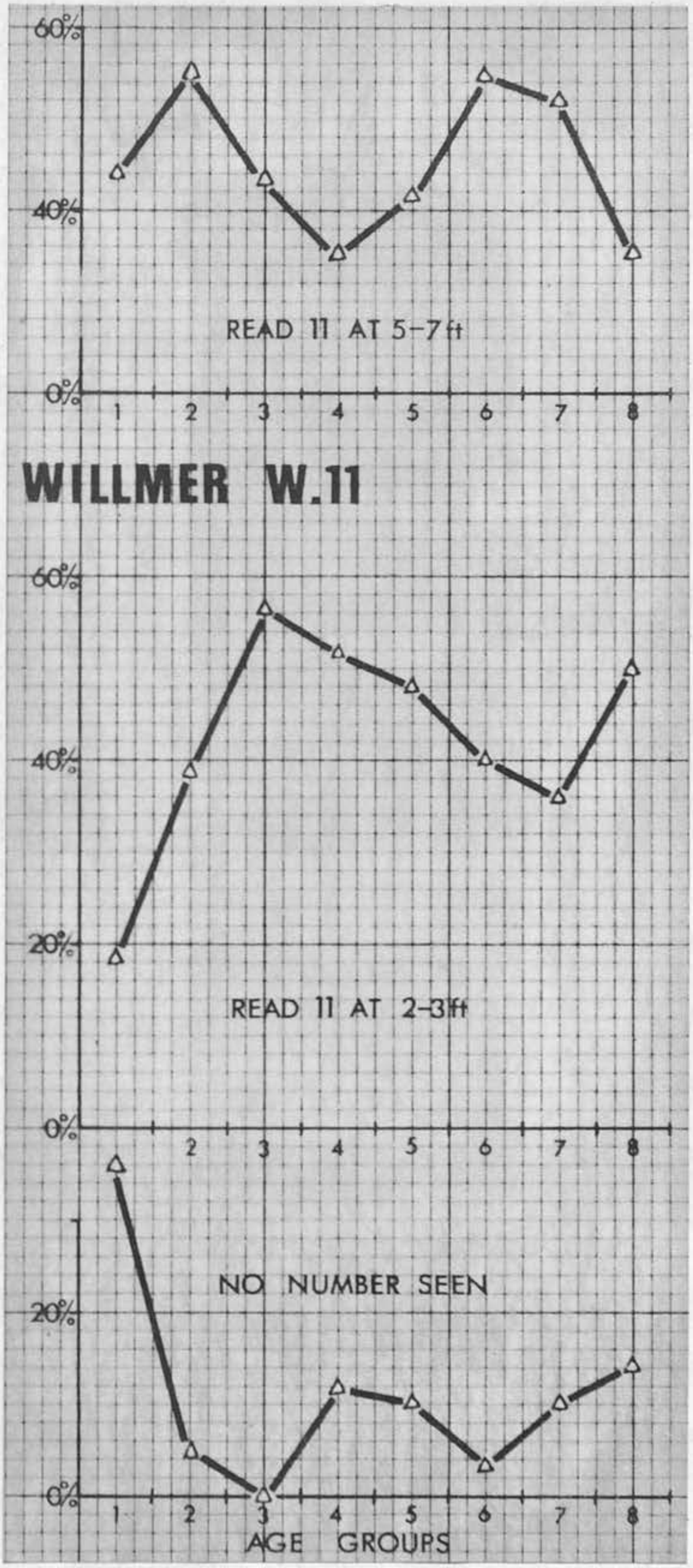

WILLMER W.2
MALE
FEMALE
BOTH
CLEAR RECOGNITION OF 2
AGE GROUPS
READ AS 8 3 5
NO RECOGNITION OF ANY NUMBER
AGE GROUPS
100%
80%
60%
40%
20%
0%
1 2 3 4 5 6 7 8

Incidentally, it is interesting to note that this plate is used in the laboratory for routine testing as a second check on red-green types of defect. All protanopes, deuteranopes, extreme deuteranomalous and protanomalous read only one of the squares, the blue one. They never see the green square.

(ii) In analysing Willmer's plate W. 2, it should be remembered that it was designed with the direct intention of detecting physiological foveal tritanopia and was found by Kalmus to be a very effective screening plate for tritan defects. There are three diagrams dealing with this plate, the first giving the incidence of those reading a clear 2, the second showing the incidence of no recognition of any number, and the third giving the incidence of those reading either 8, 3 or 5.

The two curves showing clear recognition and no recognition at all, give the familiar pattern of an inverted 'V' curve in the case of clear recognition and a 'U' curve in the case of no recognition of any number. Compared with Farnsworth's F. 5 plate the curve is much less steep. The majority of people between 15 and 50 read this plate, the percentage varying between 60% and 90% and it is only in the last two age groups that clear recognition drops to 40% and 20%. From this account it will be seen that between 50% and 70% of older people perform on these plates as if they were tritanopes.

The last of the three frequency curves shows the incidence of those who instead of reading 2, read either 3, 8 or 5. Here the incidence in all the age groups is about 8 to 10% with a slight peak for the adolescent group (18%).



Willmer's plate W. 11 gives rather erratic results as the following three diagrams will illustrate.

The first shows clear readings of 11 at a distance of 5 to 7 ft, that is, at a greater distance. Note that here there are two peaks of readings one for the second age group and one in the sixth to seventh age groups, where the incidence is 50%. In the very young, the middle aged and the very old it is read only by 40%.

The second curve shows the frequency distribution of percentages of those who read a clear 11 at a near distance, while the last curve refers to incidence of those who see no number at all, whatever the distance of viewing.

IV. Conclusions - There are very few good plates for detecting tritan types of defect and those that are in existence seem to show a variety of results. There is less consistency among them than is found among the pseudo-isochromatic plates that test red-green deficiencies.

This might be explained partly by the fact that interest in the tritan type of defect is only of recent origin and therefore as far as construction is concerned these plates have not yet reached the same level of perfection found in the red-green type of plates.

However, there are other reasons for this. Detecting deficiencies at the blue end of the spectrum proves difficult since factors other than congenital deficiencies of the receptor system have to be taken into account. The most important factor is, of course, pigmentary changes in the macula. It is relatively easy to obliterate the effect of macular absorption at 460 mu when colorimetric instruments are used in the study of colour discrimination. This is done by selecting the wavelengths for matching so that the 460 mu point is avoided. When plates with broad band reflectances are utilised it is very difficult, - even impossible - to exclude the 460 mu point from the construction of the plate. All pigments which absorb or reflect the blue end of the spectrum do so as broad band reflectance or absorption media including all wavelengths from around 380 to 500 mu. If 460 mu is excluded and only the violet part of the spectrum is left then the blues that could be utilised would be of very low luminance value and this forms a handicap in the construction of successful plates.

The results obtained from testing with the three plates, (Willmer W. 2

and Farnsworth F. 2 and F. 5) show the familiar configuration found in the other tests - a build-up of people with good discrimination up to the third or fourth age group, and then a gradual increase in the incidence of misreadings or of non-readings among the old. Few responses in this study are of the expected tritan type, but it could be inferred from Farnsworth's F. 2 plate that approximately $2\frac{1}{2}\%$ of those tested possess this type of defect. The incidence of such defects is not confined to the older people but is evenly distributed throughout the age groups.

The effect of sex upon reading of the tritan plates, though it exists, follows no particular pattern and on the whole from the third age group onwards there are no real differences. The differences between boys and girls in the incidence of misreadings or correct readings for the two youngest age groups are sometimes as high as 10 or 15%. However, no consistency is observed. Boys read the W. 2 plate correctly more readily than girls, but there is no difference in the percentages between the sexes, when non recognitions of the W. 2 plate are taken into account. The same is true of Farnsworth's F. 5 plate. Again boys have a higher incidence of correct readings for the two age groups than girls, but the analysis of 'no recognition' of the plate reveals no significant differences between the two sexes, though it should be mentioned that girls have a much higher percentage (38% against 8% for boys) of reading the F. 5 plate as '8' in the 2nd age group.

At present, because of the small numbers involved no finer analysis can be undertaken, to explain the consistent 'better' reading of the W. 2 and F. 5 plates by boys of 5 to 15.

3 : 4 COLORIMETRY AND SPECTROPHOTOMETRY

Before giving an account of the results for the anomaloscope, a section dealing with the spectrophotometric and colorimetric analysis of the pseudo-isochromatic plates is included.

In this section there is colorimetric evidence which will be useful for evaluating the effectiveness of the various tests and this should also explain why some tests (e.g. Dvorine) are more sensitive to the age variable than others - (say the Ishihara).

The sub-section explaining the 'terms' used in the science of colour, has been extended to provide definitions that will also be used when the anomaloscope is described.

(a) Definitions

By spectrophotometry we refer to the relative measurement of radiant energy or radiant flux as a function of wavelength. For visual purposes, spectrophotometry is restricted to measurements in the visible spectrum i.e. approximately from 380 mu to 750 mu. The term relative in this connection, refers to the fact that in spectrophotometry the measurements are always made relative to some standard and the standard used here was magnesium oxide. In this way what is measured is merely the ratio of two quantities wavelength by wavelength throughout the spectrum range. It has become axiomatic in psycho-physical colorimetry that spectrophotometry is the fundamental basis for standardisation and specification, and it is the only fundamental means of analysing a colour for research or other purposes. This is the way to measure, independent of deviations of colour vision existing even among normal observers, and, thus, it is the basis for colour specification. But a spectrophotometric curve is quite inadequate as a colour specification in itself. In order to provide adequate specifications with tolerances, the spectrophotometric data must be converted to some form of tristimulus specification by means of methods and data, defining a standard observer, a standard co-ordinate system, and a standard illuminant. In common practice colour is regarded as the property of objects but we have to re-define this to include that it is also the property of the light source as well as the object. Some even go so far as to say that it is a property of light alone (Judd 1950, Gibson 1949). However, any example of metameric pairs of stimuli will immediately bring to mind the fact that there is also the property of the observer to be considered. Thus, when specifying colour we must not only include the spectral composition of the radiant energy reaching the eye of the observer, but also the properties of the observer as such. If a normal observer attempts to do matching, be it with skeins or paints, or lights, or pigments, or mixtures on a sector disc, he will find that he requires three primaries. In the case of the paints used in primary schools, red, yellow and blue are the required primaries, in photography and lithography the colours are magenta, yellow, cyan, while with lights they are red, green and blue. Three primaries are the irreducible minimum and thus it is concluded that normal colour vision is tri-dimensional. Because of this, colour specification is or can be expressed by three numbers.

In the example given (i. e. paints, spotlight, sector disc, etc.) by adjustment of three variables, the observer obtained a colour match, that is he has set up a second stimulus equivalent to the first. The mixture is equivalent to the unknown in colour, but not in spectral composition, and the unknown and the mixture therefore form a metameric pair. Thus though spectrophotometrically they may differ very much they manage to have exactly the same appearance to a normal observer. Such equivalent stimuli are called metamers and are said to form a metameric pair. The properties of the observer are completely defined by stating which pair of stimuli are found by him to be metameric, and modern colorimetry is based upon spectrophotometry interpreted according to the properties of an observer.

From a knowledge of spectral metamers it has been possible to summarise concisely the properties of an average normal eye and this has been made in accordance with the principle known as Grassman's Law which was foreshadowed by Newton's Law of colour mixture :-

This states that if a light composed of known amounts of three components called primaries, is equivalent in colour to an unknown light, the three known amounts must be used as a colour specification for this light. These amounts are called the Stimulus Values of the colour. Grassman's Law states that when equivalent lights are added to equivalent lights the sums are equivalent.

Thus, if an unknown spot of colour were matched by shining on a particular spot on a white screen, two component spotlights of tristimulus values $X_1 Y_1 Z_1$ and $X_2 Y_2 Z_2$ respectively, by Grassman's Law the tristimulus values of X Y Z of the unknown spot of colour would be simply -

$$X = X_1 + X_2$$

$$Y = Y_1 + Y_2$$

$$Z = Z_1 + Z_2$$

The X Y Z values of the primaries required to produce the stimulus equivalent to any light beam may therefore be found by adding together the tristimulus values of the various parts of its spectrum. Modern colorimetry is based on this application of Grassman's Law.

All colour data that are used in this thesis are expressed in terms of the tristimulus system accepted in 1931 by the International Commission on Illumination (C. I. E.) for foveal viewing. The three primaries used in this system are imaginary primaries expressed as X Y Z and these completely specify the various mathematical functions of the system. Thus specification of a colour involves finding the relative amounts of the threeprimaries required to match the colour in question under a given illuminant. The tristimulus values of spectrum stimuli of unit irradiance have been found empirically and these are usually given in text-books dealing with the C. I. E. system (e. g. Le Grand 1958 etc.).

In practical colorimetry, the tristimulus values X Y Z for any stimulus can be calculated from the following formula and these were used for the determination of the various stimuli used in this research.

(i) For 'primary sources'

$$X = \sum \bar{x}\, E_\lambda \Delta\lambda$$

$$Y = \sum \bar{y}\, E_\lambda \Delta\lambda$$

$$Z = \sum \bar{z}\, E_\lambda \Delta\lambda$$

Where (a) $\sum$ denotes summation over all equal intervals of

(b) The actual size — for this type of calculation was 5 mμ

(c) E_λ is the <u>relative</u> spectral emittance of the source,

(d) $\bar{x}$, $\bar{y}$, $\bar{z}$, are the tristimulus values of the spectrum colours of the standard observer.

(ii) For 'secondary sources' (i. e. primary source plus filter)

$$X = \sum \bar{x}\, E_\lambda T_\lambda \Delta\lambda$$

$$Y = \sum \bar{y}\, E_\lambda T_\lambda \Delta\lambda$$

$$Z = \sum \bar{z}\, E_\lambda T_\lambda \Delta\lambda$$

Where (a) T_λ is the 'transmission factor' of the filter for wavelength

(b) $\Delta\lambda$ was 10 mu.

(iii) For 'duffising materials' (i. e. surface objects)

$$X = \sum \bar{x}\, E_\lambda \rho_\lambda \Delta\lambda$$

$$Y = \sum \bar{y}\, E_\lambda \rho_\lambda \Delta\lambda$$

$$Z = \sum \bar{z}\, E_\lambda \rho_\lambda \Delta\lambda$$

Where (a) ρ_λ is the 'spectral reflectance factor'

(b) $\Delta\lambda$ was about 16 mu

It must be added here that the $\bar{y}$ tristimulus value also carries the luminosity function, and it is customary to express the Y value of a luminous area as its luminous or photometric brightness.

However, a more complicated concept is employed in the C.I.E. system, that is 'luminous directional reflectance', where we refer only to that flux which is reflected in the direction of the observer. This is more closely related to the actual appearance of the specimen, especially when Y is evaluated for opaque specimens. This directionality for the C.I.E. system is defined as $0/45^{o}$ viewing. All measurements in this thesis were made under such conditions both on the Tintometer and on the Spectromat. The C.I.E. specifications yield further information used in this study, namely the 'chromaticity co-ordinates'. These co-ordinates, denoted by symbols x, y, z locate colour on the C.I.E. chromaticity diagram, (known as the Maxwell triangle). This is also called a mixture diagram because it indicates, in a very simple way, the chromaticity of the colour resulting from the additive combinations of any two lights.

Chromaticity co-ordinates are computed directly from the tristimulus values by converting these into proportions thus -

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

Since the sum of the chromaticity co-ordinates is unity ($x + y + z = 1$) a knowledge of any two of them leads to the value of the third. Two types of graphs are in use, the first utilizing x and y co-ordinates, and the second utilizing x and z co-ordinates to plot the colour space.

The classical C.I.E. chromaticity diagrams used here are expressed in x and y co-ordinates, while diagrams referring to the Uniform Chromaticity Scale (U.C.S.) according to Judd (1935), use the x and z co-ordinates.

A third set of specifications are used in describing the colours of the stimuli studied in colour vision tests used in this research, namely

'dominant wavelength' - λ D

and 'excitation purity' - pe

Instead of specifying the colour by its co-ordinates x and y on the C.I.E. diagram, the dominant wavelength (λD) and its excitation purity (pe) are sometimes given, as this is more meaningful to those who have some idea of the hue from the value of the wavelength, and of saturation from the ratio of purity.

<u>Dominant wavelength</u> - This is determined by drawing a line within the C.I.E. chromaticity diagram between the chromaticity point of a reference source and the chromaticity point of a sample, and extending the line until it intersects the spectrum locus. The dominant wavelength is that wavelength at which this line intersects the spectrum locus.

Whenever the value of a dominant wavelength is quoted in this thesis the chromaticity point of the illuminant 'C' is taken as a reference point that is :

$$x = 0.31006$$

$$y = 0.31616$$

In practice, the choice of a particular 'W' is either of one under which the matching was made, or of some other 'W' chosen arbitrarily, as for example equi-energy source 'E'. As we have used some four different sources, three of which lie very near the location of illuminant 'C', it was decided that the dominant wavelength notations for the anomaloscope arbitrary units should also be referred to this source, irrespective of the fact that the real illuminant in the anomaloscope has a colour temperature different from that of illuminant 'C'.

In this way the hues, as found from the colorimetric analysis of the various tests can be compared with each other. In the case of λD for the anomaloscope units, other factors also indicated the use of illuminant 'C' as the achromatic reference point. The chromaticity co-ordinates for the anomaloscope source lie near the loci of the middle part of the red-green and yellow-blue equations, and in such cases as this the concept of λD has very little meaning as it introduces a large degree of error in the graphical calculations necessary to find this value. Therefore a position of a source further away from the R-G and Y-B line was chosen.

For colours that fall into the region of non-spectral colours (i.e. the region of purple) the 'complementary wavelength' is determined because no dominant wavelength exists for such colours. To denote complementary wavelength, two conventions are used, - either a minus sign is placed in front of the wavelength value e.g. -530 mu or the letter 'c' is placed after this value e.g. 530c mu. The latter convention is used in this work.

Excitation purity (pe) is the ratio of the distance between the chromaticity point of the reference source and that of the sample, to the distance between the chromaticity point of the source and the point on the spectrum locus representing the dominant wavelength of the sample.

Pe is computed from the following equation :-

$$pe = \frac{x - xa}{xb - xa} = \frac{y - ya}{yb - ya}$$

Where pe represents the excitation purity, x, y define the chromaticity point of the sample, xa, ya define the chromaticity point of the light source, and xb, yb define the chromaticity point on the spectrum locus or the non-spectral boundary.

Lastly, in some instances, overall colour differences will be given. These are quantitative values, which can be obtained and calculated from the C.I.E. data described so far.

There are many colour difference formulas, all of which are valid only to the extent that the special models on which they are based are valid representations of uniform colour space. Unfortunately, so far no one has succeeded in producing such a uniform space. However, the formulae in use are at least a fair approximation to the ideally uniform colour space.

There are two- and three-dimensional uniform chromaticity scales. In the first type, hue and saturation are taken together, and colour difference is calculated from this. This approach was used when finding colour differences for the Pickford anomaloscope. These results are a reasonable approximation to the ideal values since the 'lightness' variable is not important in this case, as matching is always done when the luminosities of the two fields are equated.

In all other calculations of colour difference, a tri-dimensional colour scale is used, where in addition to hue and saturation, a 'lightness' scale is taken into consideration. The colour difference formula used here is that of Nickerson and Stultz (1942), which is based on Adams' chromatic-value diagram (Adams 1942), combined with the Munsell value scale. This is the most widely used method of calculating colour differences between two samples whose tristimulus values X Y Z are known by spectrophotometric measurements The formula reads as follows :-

$$\Delta E = \left\{ (0.23\,\Delta V_y)^2 + \left(\Delta(V_x - V_y)\right)^2 + \left(0.4\,\Delta(V_z - V_y)\right)^2 \right\}^{\frac{1}{2}}$$

where E = colour difference

V_y = the Munsell value function,

V_x = the Munsell value found from the same function by setting $Y = X_c$

V_z = the Munsell value for $Y = Z_c$

The Munsell value functions V_x, V_y and V_z were obtained from a paper by Nickerson (1950).

Colour differences are expressed in N. B. S. (National Bureau of Standards) units. One N. B. S. unit is equivalent to

0.10 Munsell value step

0.15 Munsell chrome step

2.5 Munsell hue step at chrome /1.

In psycho-physical terminology, the N. B. S. unit is about five times greater than the smallest perceptible difference under the best experimental conditions.

Differences of one N. B. S. unit or less are usually disregarded in commercial transactions.

Perhaps a last word should be added to give an indication of the reliability of the photometric and colorimetric data obtained here. The question of reliability of measurement using the Pretema Spectromat has been fully discussed by G. Robertson (B. Sc thesis 1964). His conclusions were, that in measurements involving 5 mm apertures (in the case of pseudo-isochromatic tests) the accuracy is of the order :-

maximum $\Delta x = 0.005$ minimum $\Delta x = 0.013$

$\Delta y = 0.006$ $\Delta y = 0.010$

Since this study has been concluded, a method overcoming the loss of luminosity for 5 mm apertures has been devised so that it is possible now to make further photometric readings utilising the full 100% extent of the scale on the oscilloscope and this has considerably increased the accuracy of the measurements. In measurements involving apertures larger than 5 mm (C. A. T., tritan plates, etc.) the expected accuracy is even higher.

In terms of colour difference measurements, differences of 1 to 1.5 N. B. S. begin to be meaningful and significant while smaller differences might be due to errors inherent in the measuring techniques employed on the Spectromat.

For tests described here photometric data giving reflectance or transmission values are represented by curves on the appropriate graphs.

Colorimetric data include tristimulus values, x and y co-ordinates, luminous value Y, or, if this relative measure is inappropriate, physical photometric quantities of luminance or illumination are given. Excitation purity and the dominant wavelength are also quoted.

In the case of tests consisting of object colours (for example coloured surfaces on pseudo-isochromatic plates) the above data refer to the individual colour dots of which the test is composed. Of course, such an approach is highly analytical, for in the last instance we really do not know whether the 'eye' sees the dots individually and then summates them to make the total 'figure' against the 'background' or whether this analytical process is continued right up to the central cortical areas of the optical lobes, where 'inferences' are made. If the visual process is one of comparing the totality of the colour impression of the figure-background configuration, then the values given for the individual colour dots will not be exactly the same as given here, because the effect of temporal and spatial summation would alter the actual colour perceived. However, even in such a case, these additional complications of the visual system will not affect the positive contribution the analytical approach can offer. In the 'normal' eye this could be looked upon as a factor having a constant value, thus the relative positions in the colour space would not be altered. In the 'colour defective' eye these factors will be different for each different type of 'deviant visual system' and, although the relationships found by this means will not always hold good, this would be true irrespective of whichever analysis of the notation of the apparent colours had been made.

Photographic evidence - Whenever possible photographic evidence of the relative luminance values of the individual dots, or of the figure-background configuration for the pseudo-isochromatic plates is included. This is done to clarify and confirm the results obtained from photometric analysis of the plates. This evidence is important and will be used to illustrate the nature and complexity of the construction of such tests and the complications such factors introduce when any analysis of performances is made.

One might ask why photometric and colorimetric analyses are necessary. In the various research programmes that have been initiated since the last war to evaluate the usefulness of pseudo-isochromatic tests, it was usual to compare scores for a group of subjects on the plates with some measure such as anomaloscope data for the same group of subjects. As the diagnostic efficiency of the screening of normal versus colour vision defective has been taken as the main criterion for the validity of using such tests, these tests have been compared and tables constructed to show which were better. The Dvorine and Ishihara tests were constructed on purely empirical grounds, and the colours used in their construction were selected on a trial and error basis, only those plates being retained which were shown to be effective in diagnosing colour vision defects. In this research the aim is different. The aim is to give an objective analysis of what these tests are actually testing. In this way it is possible to ascertain with a high degree of accuracy, in terms of physics and psycho-physics, the nature of the stimulus that each particular test and each particular plate in the test presents to the subject. Secondly, this allows us to compare the tests with each other in an objective manner. It should, for example, indicate why some tests pass and others fail a certain type of subject. At the same time an attempt is to be made to find how colorimetric analysis can help us to evaluate certain theoretical postulates currently in use by people studying colour vision, namely the confusion line theory and the so-called isochromatic lines of Pitt (1935).

(b) Measurements

The effect of age on the performance on the tests used varies greatly. The Dvoriné shows the greatest variation in scores for the different age groups, while the Ishihara seems to be less effected. Even the Tritan plates are affected to some extent, registering difference in readings for both young and older subjects, (e.g. Farnsworth F.5 plate) while yet others such as Willmer's W.11 plate shows no consistent results in this respect. In the historical section of this thesis the CAT test was found to be sensitive to age variations, and, in the section on Age and Colour Vision after 1958, we shall see that the 100 Hue Test is again a sensitive measure.

A rather perplexing problem now faces us. It might be said that the difference between finding an ageing effect and not finding it was one of differences in administering the tests or in analysing the results of the various test plates. This has already been discussed in the historical section where, though this was conceded to be possible, it was also pointed out that other factors play an important role. It should be mentioned for example that, in this research, all the pseudo-isochromatic tests were administered in an identical manner and they were used on the same population. Yet in spite of this differences were still observed. So another explanation is necessary. The differences in the structure of these tests is perhaps one of the most important factors giving rise to such results. Photometric and colorimetric analysis of the tests were made as these are the only means by which we can demonstrate any differences in the construction of the plates, which might affect the ease or difficulty with which each plate was read.

The pattern of the analysis of tests will be repeated for each of the tests in turn. Photometric data, consisting of reflectance curves is given. Such information is valuable as it is not only the basis for further colorimetric calculations, but also gives some indication of the type of pigments used in the construction of the coloured dots in a given test plate. This is an important point, not often stressed by those who make such tests, and scarcely appreciated by those who used them. Pigments with 'ideal' absorption curves, lead to more stable colours - and this recognition or lack of it will depend on the degree and type of defect of a given subject. If the coloured dot represents a stimulus with a peak at 600 mm., and without any secondary peaks, say, at 430 or 460 mm., then such a stimulus - though evoking the sensation of red, will not be detected by those in whom the red end of the spectrum is shortened. However, a similar 'red' with a secondary peak at the blue end of the spectrum will be seen as a 'purplish' or even 'bluish' colour by such defective people. Thus a stimulus intended to evoke the sensation of 'red' and nothing else, will, - in the case of a stimulus with two peaks - give rise to a 'secondary' sensation, as already described.

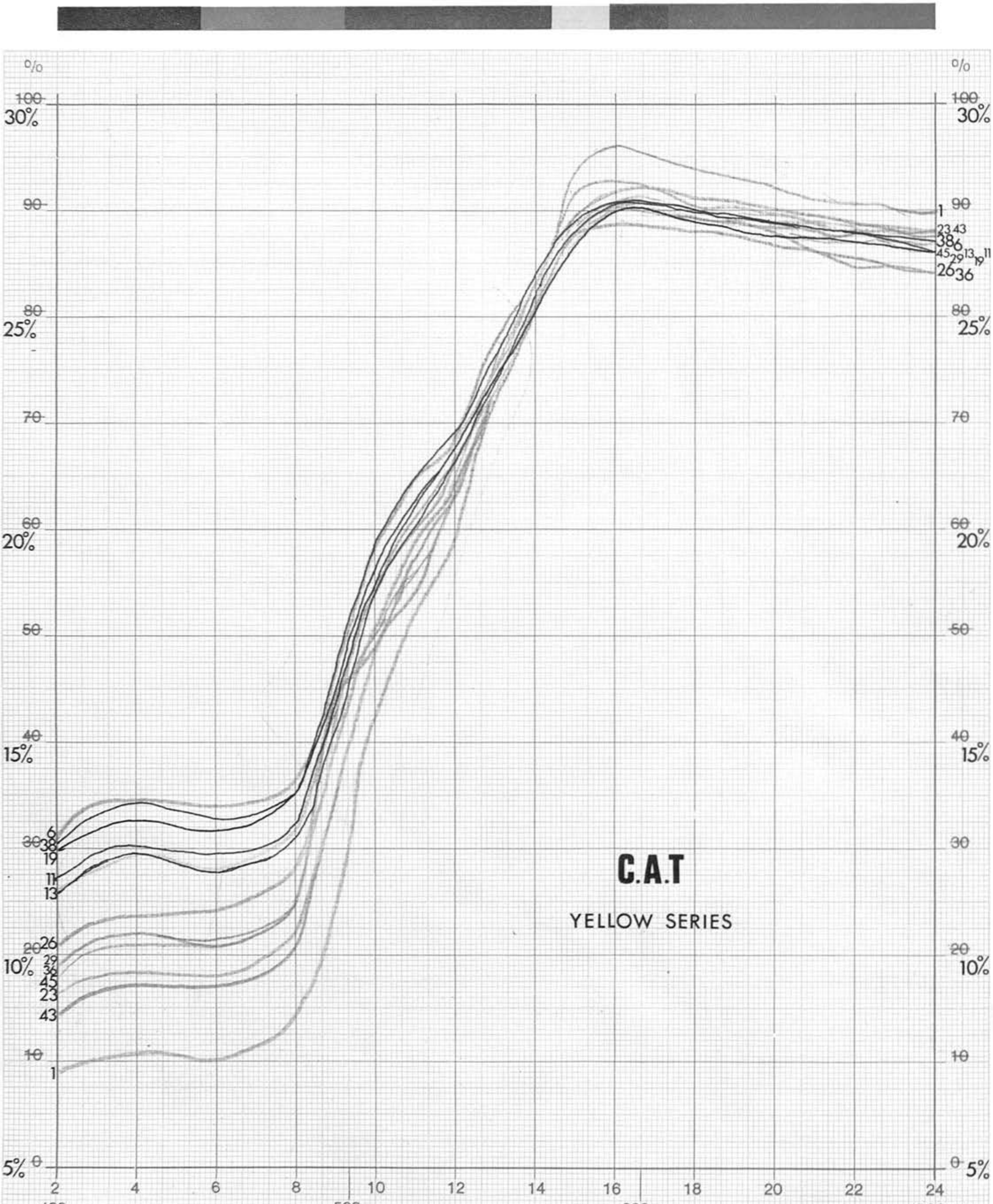
The photometric curves also allow us to say how many actual 'colours' have been used in constructing the test. Bearing in mind the problem of metamerism, it will be immediately obvious that it is not enough to say that only 3 or 4 colours were used if for example any one of the four is in a metameric relationship with any of the other colours mentioned. Subjects with even a slight deviation of 'colour balance' from the 'standard observer' may perceive more than just 3 or 4 colours, for they will have different metameric relationships from those employed in the test.

Colorimetric analysis will be used to study the overall position of colours employed in the construction of all the plates. A "bird's eye view" of the general position of the test within the colour space will thus be obtained. Questions such as, which is the most predominant axis for the colour combinations and whether this is only in the red-green direction, or also contains elements of blue and yellow, can now be answered. In other words, it is now possible to say, whether the tests test only red-green confusions, or whether they unwittingly involve yellow-blue discrimination also.

An analysis is made of those individual plates, which are of particular interest. Some plates (e.g. in the Ishihara) are discussed at length because it is possible now to explain and illustrate why this particular plate was confused so often.

Lastly, where it is of interest, measurements of the differences between the colour position of the dots of the figure and background are expressed in terms of N.B.S. units showing the total colour difference present.

I. CAT and 100-Hue test. - Earlier in the historical section it was noted that the CAT test was used by some research workers, and that it was the only test sensitive enough to measure differences in performance over a large number of age groups. For this reason data for the objective measurements of this test is now quoted along with similar data for the 100-Hue test which is included because it was the main test in Verriest's (1962) study which is to be discussed



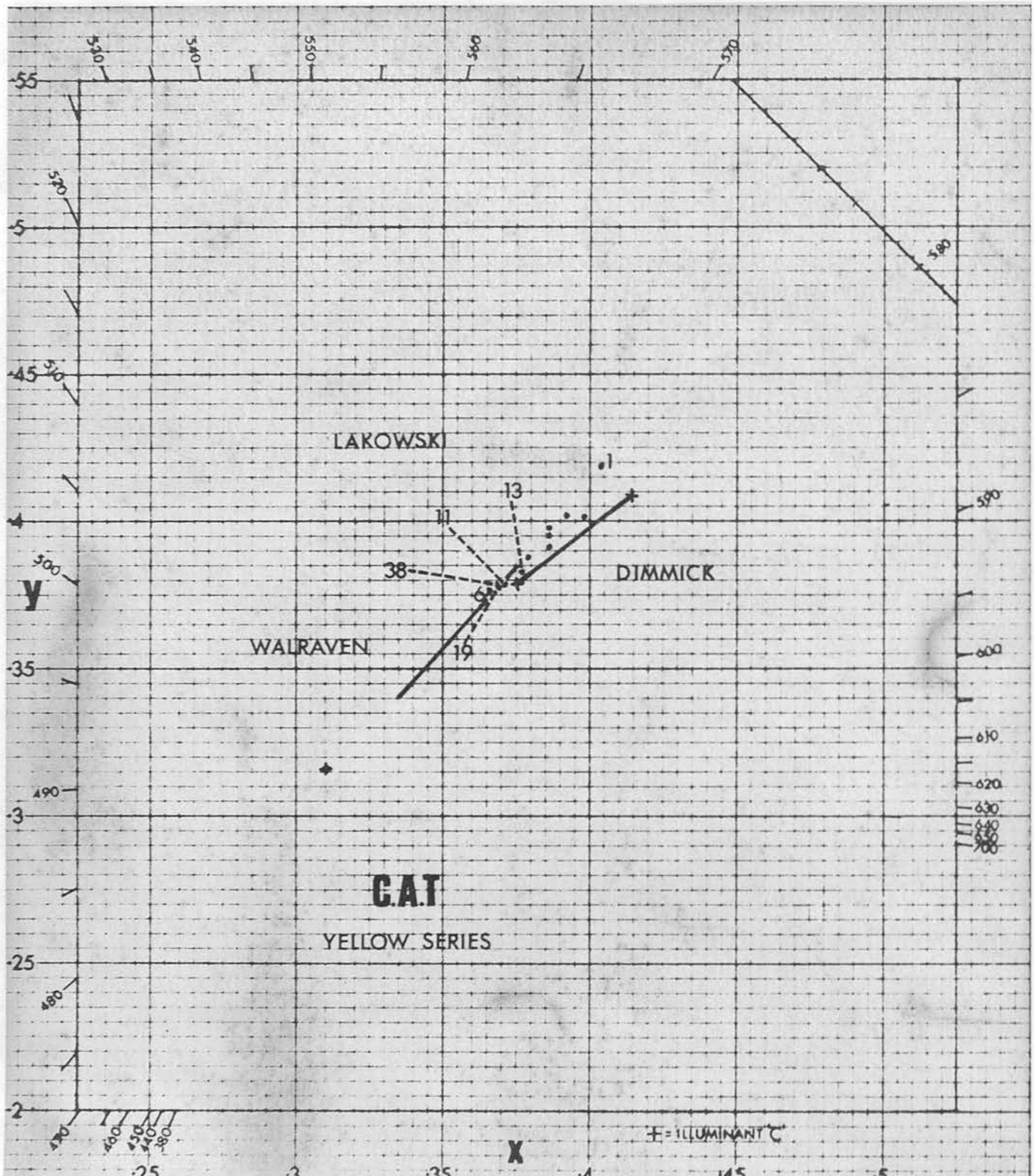
C.A.T
YELLOW SERIES
%
100
90
80
70
60
50
40
30
20
10
0
30%
25%
20%
15%
10%
5%
2
4
6
8
10
12
14
16
18
20
22
24
400
500
600
700 mμ

later. The data presented should convince us of why it is that these tests can be instrumental in detecting minute losses of color discrimination, and why they can indicate the variations in performance which relate to the different age groups.

(i) The CAT test - Dimmick (1956) did not give direct colorimetric data obtained from the measurement of the individual chips of the CAT test, but rather quoted the co-ordinates for theequivalent Munsell colours representing the four series of the test. The data shown here was obtained from the direct measurements of the various chips. Colorimetric and photometric results were obtained on the Spectromat using $0/45^{\circ}$ geometry and magnesium oxide as the standard. Measurements were confined to one series only - yellow.

First let us examine the reflectance curves. If we look at the inclination of the individual curves of the series it will be seen that there are only slight variations between each of them. (The position of the lower and the upper ends of the slopes in relation to the wavelength axis determines the dominant wavelength for these chips). There are variations at the 'foot' of this curve, and this is what determines the variations in saturation among the individual chips. In the plateau at the top of the graph the individual curves are tightly packed together which indicates that there should be little variation in brightness between the chips.

Judging from the photometric curves, we can expect very little variation in hue in this series but there should be some variations in saturation and rather less in brightness. The C.I.E. diagram with the loci of the



LAKOWSKI
DIMMICK
WALRAVEN
C.A.T
YELLOW SERIES
y
x
+ = ILLUMINANT 'C'
·55
·5
·45
·4
·35
·3
·25
·2
·25
·3
·35
·4
·45
·5

individual chips of this series, shows that they form a line in relation to the locus of illuminant 'C' and the spectrum locus, and this points to a dominant wavelength between 576 and 577 mu indicating very little variation of hue. The greatest variation lies in the changes in saturation which can be seen from the movement of the loci for that particular chip from the vicinity of the spectrum locus line towards the centre of the diagram. Besides the individual positions for the chips measured the diagram also shows 'two lines' representing the positions of this series given by Dimmick (1956) and measured by Walraven et al (1956). There are discrepancies in the three accounts, but the greatest discrepancy is between Walraven's data and that of Dimmick and Lakowski. The differences between the two last authors could be due to using different geometries for the illuminants used in the measurements. In the Spectromat directional illumination was used while Dimmick used Munsell data as obtained from measurements under diffused illumination.

Colour discrimination has to be very fine before high scores can be obtained, on this test. A calculation in terms of total colour difference was made for one of the chips on the test board, namely number 11. Four chips from the dispenser are accepted as correct solutions, but, a weighted score is used.

The colour differences calculated in N.B.S. units

of board Numer 11 with chip 11 = 0.84

11 with chip 13 = 1.00

11 with chip 38 = 2.5

11 with chip 19 = 2.6

The number of N.B.S. units represents the total 'colour difference' in the colour space for the standard observer between a given chip (No. 11) and the chips matched with it.

It is interesting to note that the placing given in the manual is not consistent with our results :-

Chip number	11	13	19	38
Weighted score	3	2	2	1

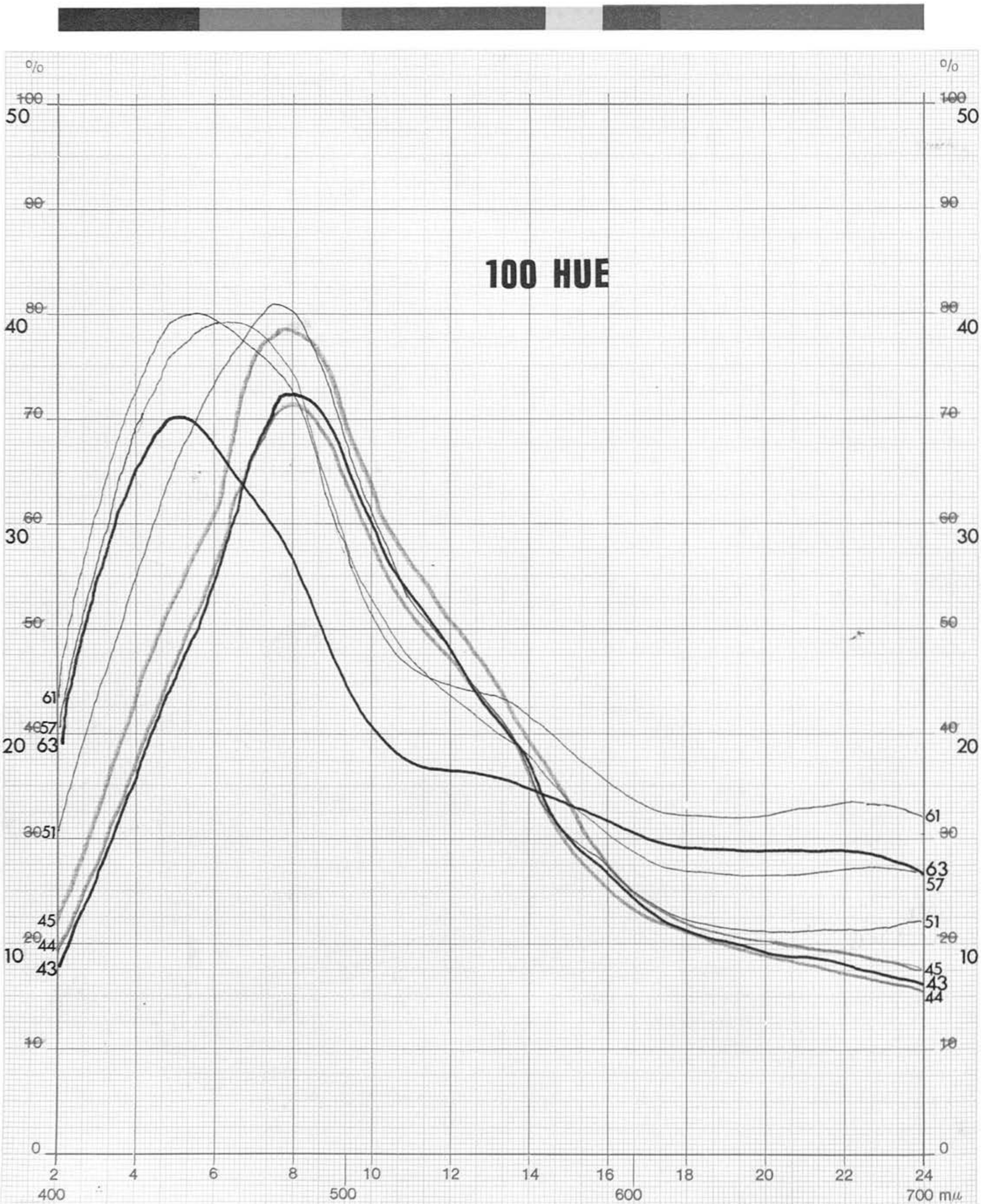
Chips 13 - 19 are allowed an equal score, but in terms of N.B.S. units, 19 is almost $2\frac{1}{2}$ times further away from board 11 than chip 13.

This is rather an odd weighting - especially as there seems to be no justification for it from the empirical studies on the order of frequency for the selection of a given chip and a given 'board' position. (See the following table).

Author	Order of Frequency			
	1	2	3	4
Dimmick (1956)	11	13	38	19
Adams (1958)	11	19	13	38
Walraven (1956)	13	11	38	19

The other series of this test have not been measured, but it is safe to assume that color differences will not be greater than 5 N.B.S. units, if a score is to be obtained. These chips which approach the board number in colour specification are likely to yield the highest score.

Thus, the CAT test presents very small differences in discrimination

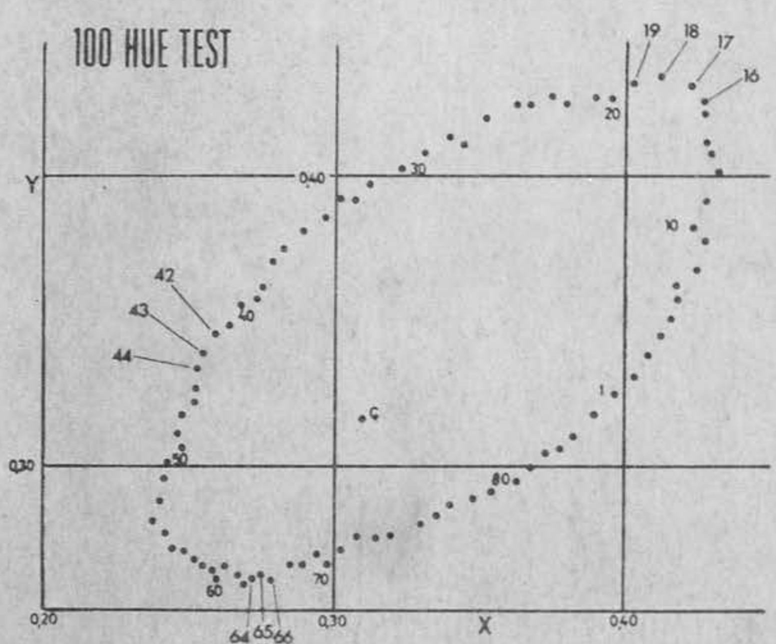

100 HUE
%
50
40
30
20
10
0
61
57
63
51
45
44
43
2
4
6
8
10
12
14
16
18
20
22
24
400
500
600
700 mμ

and as such its results for the age population can only indicate whether there is a change in the frequency of those with colour discrimination of between 1 to 3 N.B.S. units. The test cannot indicate the extent of the loss of colour discrimination in an age population.

(ii) Farnsworth Munsell 100-Hue Test - From visual inspection of this test it can be seen that the colour steps of a series are very close together. Let us look first at the photometric curves. The graphs show representative reflectance curves of caps within the 43-63 series of the test. For the sake of clarity the number of curves has been limited. In general it will be noticed that the individual curves have only one sharp peak in the visible spectrum at about 480 mu in this series. There must be variations in hue as the peaks vary from approximately 495 mu to 450 mu. These vary in height, and thus indicate that there are probable variations in subjective brightness between the caps of this series; The existence of differences between slopes of the curve at the red end of the spectrum also indicate the presence of saturation differences.

These are mostly very well defined curves resembling the 'bright' spectral filter characteristics of the Ilford or Kodak series, and unlike the CAT, hue changes here are quite a dominant feature of the series.

The numerals at the extremes of the spectrum for each curve on the graph represent the cap number or Munsell colour number. The two black lines mark the end points of the series, while the two purple lines representing the two adjacent caps (44-45)(along with the black line for 43) indicate the differences between adjacent members of a series.



100 HUE TEST
Y
X
0,20
0,30
0,40
0,30
0,40
C
1
10
16
17
18
19
20
30
40
42
43
44
50
60
64
65
66
70
80

The C.I.E. chromaticity diagram shows the loci for all the caps used in the 100-Hue test, and measured on the spectromat (note specially the positions for cap 43 and 63). Confirmation for the observations made from the photometric curves is obtained from this diagram where variations in saturation, and in dominant wavelength can be seen.

To show how difficult the test is for non-defective subjects the colour differences for certain selected series of caps are quoted.

Taking, for example, caps numbered 42, 43 and 44 colour difference steps of between 2.7 and 1.3 N.B.S. units are found between caps 43 and 42 and 43 and 44. Colour difference steps were also measured at another point between caps 64, 65, 66 - and again the differences obtained were 1.8 N.B.S. units between 64 and 65, and 2.2 N.B.S. units between 65 and 66. Further samples were measured for the red part of the colour circle between caps 16, 17, 18, 19, and here the colour differences seemed to be larger than in the other two areas - 16 to 17 - 4.9 N.B.S.

17 to 18 - 2.3 N.B.S.

18 to 19 - 6.6 N.B.S. units

The mean for the seven measurements is approximately 3.1 N.B.S. units and the scatter is from as little as 1.3 to 6.6 N.B.S. units. Therefore the test can show slightly greater colour differences than were found in the CAT, but because of the way in which the testee is required to take the test (namely, step by step in a 'colour series' for each of the series of the test) it is capable of testing almost continuously where and to what extent a subject has difficulty in discriminating colour. It can detect, on the one hand, those who make only small mismatchings of between 1.5 and 2 N.B.S. units, and on the other hand, it can measure the extent of very gross mismatchings (say 10 or 15 caps) showing colour differences of between 11 - 20 N.B.S. units. (It is this type of difference that pseudo-isochromatic plates are built on). Thus the 100-Hue test is capable of detecting both very minute and very great losses.

This should be remembered when the results of Verriest's findings in his study of age and colour vision are discussed. It is also possible now to give a more precise indication of what the colour difficulties were that faced the subjects in Smith's study. If only every second Munsell step was used, the colour differences were between 3 to 8 N.B.S. units.

II. Ishihara and Dvorine tests - To allow us to compare the features which they have in common, and also illustrate more forcibly where they differ, the Ishihara and Dvorine tests are discussed together. Where either test has some unique feature, this will be examined separately. It is hoped that in the long run, this method will serve to explain more pungently how such pseudo-isochromatic tests are constructed than mere cataloguing of the separate plates for each test would do.

In terms of the construction of the pseudo-isochromatic tests, there are four distinct types of plates used :

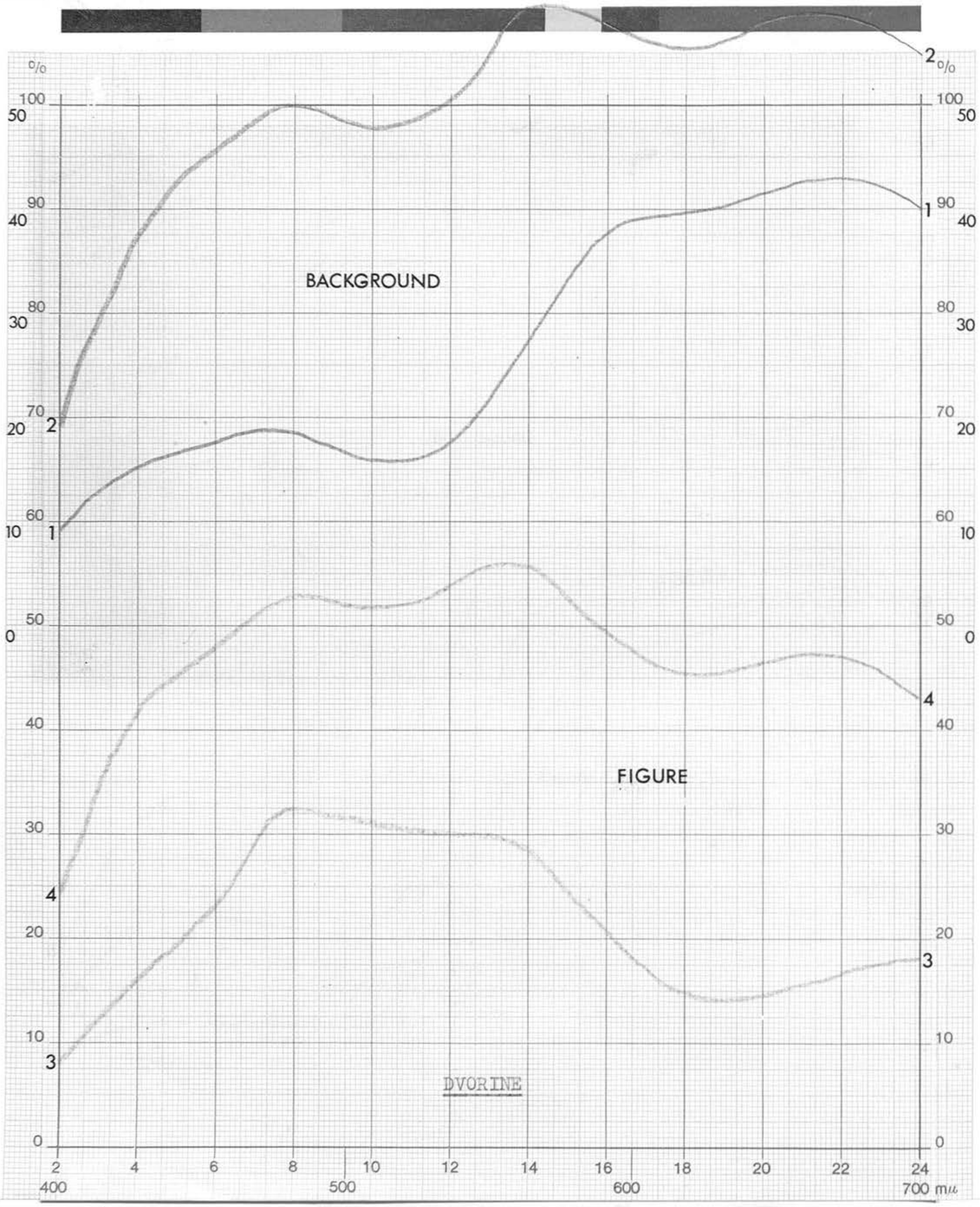
1) the transformation type of plate;

2) the vanishing digit type;

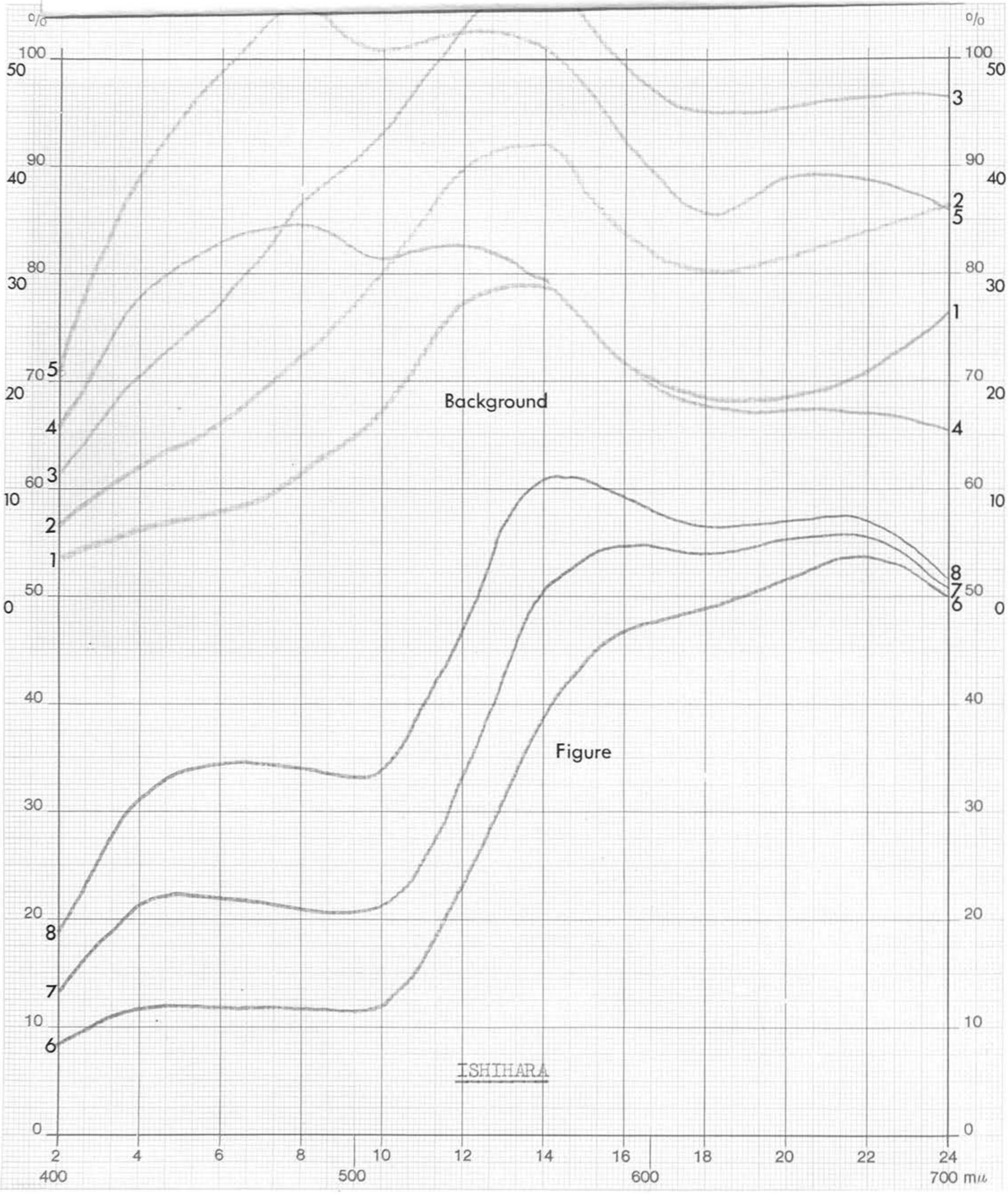
3) the hidden digit type;

and 4) the qualitative type.

All four types of plate are employed in the Ishihara, whereas the



BACKGROUND
FIGURE
DVORINE
%
100
90
80
70
60
50
40
30
20
10
0
50
40
30
20
10
0
1
2
3
4
2
4
6
8
10
12
14
16
18
20
22
24
400
500
600
700 mμ



%
100
90
80
70
60
50
40
30
20
10
0
50
40
30
20
10
0
1
2
3
4
5
6
7
8
Background
Figure
ISHIHARA
2
4
6
8
10
12
14
16
18
20
22
24
400
500
600
700 mμ

Dvorine is composed mainly of types 2 and 4. The vanishing digit type is the simplest, while the transformation and hidden digit types are more complex in construction.

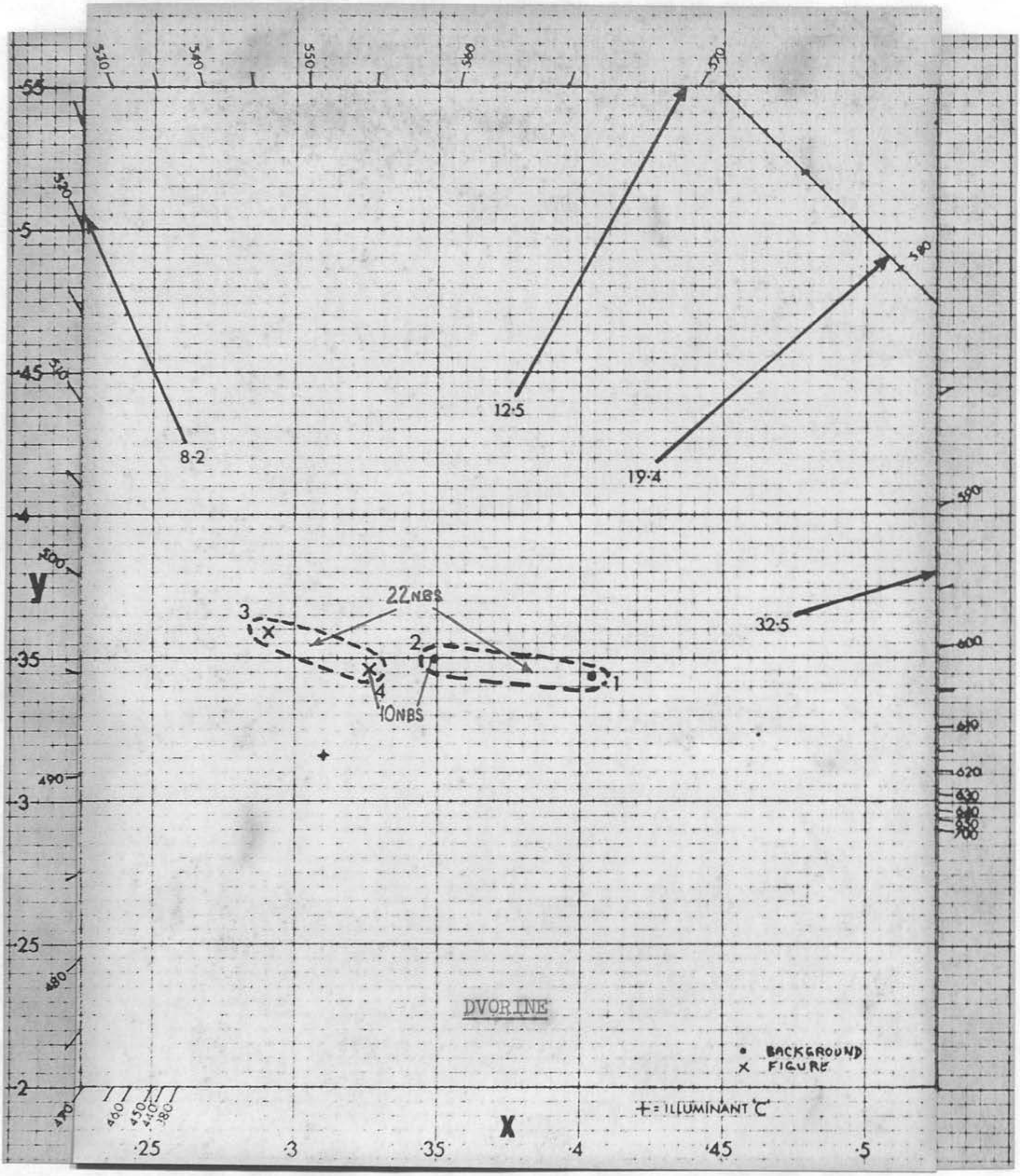
As many as twelve of the fourteen Dvorine plates and eight of the Ishihara plates are of the vanishing digit type, so this category will be analysed first.

In the Dvorine test, only four distinct colours are employed to construct these plates, while in the Ishihara as many as eight distinct colours can be identified. Photometric analysis of these plates for each test revealed a difference between the tests in the type of pigments used for constructing the plates. The curves for these particular Dvorine plates are smoother than those of the comparable Ishihara plates where they tend to more definite peaks at certain parts of the spectrum. By inference, the greens are more dominant, and the pigments used for the red end of the spectrum would arouse a clear sensation of pink or red. The curves for the Ishihara are obviously of the minus blue type, whereas in the Dvorine this is less marked, and the general level of reflectance is much lower, giving rise to a sensation of brown rather than pink or orange. The curves of the green employed in the Dvorine show a plateau at the region at maximum reflectance.

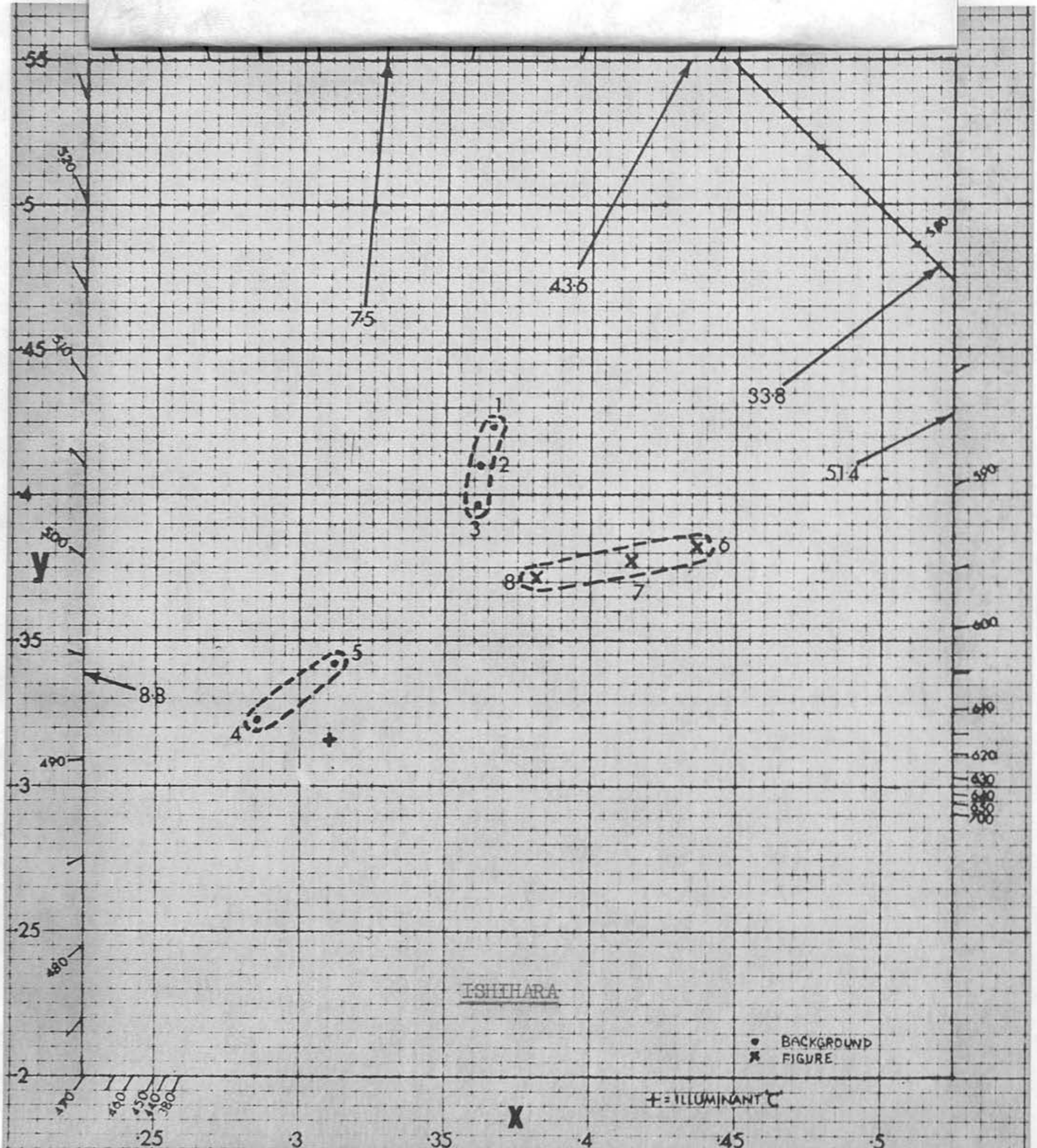
There is another essential difference between the construction of vanishing digit plates in the Dvorine and in the Ishihara.

In the Dvorine, only two hues are used for the figure, and two for the ground, and these are so chosen that both hue and colour position in the colour

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
FIGURE							
3	15:26	520.5	8.2	.291	.359	47.7	
4	14:26	569.5	12.5	.326	.346	62.3	
Bakground							
1	14:22	594	32.5	.404	.344	33.7	
2	11:24	579	19.4	.349	.350	65.2	
Table No. _ _	COLORIMETRIC DATA FOR DVORINE Series I						



y
x
·25
·3
·35
·4
·45
·5
·55
·5
·45
·4
·35
·3
·25
·2
8·2
12·5
19·4
32·5
22NBS
10NBS
1
2
3
4
DVORINE
• BACKGROUND
X FIGURE
+ = ILLUMINANT 'C'



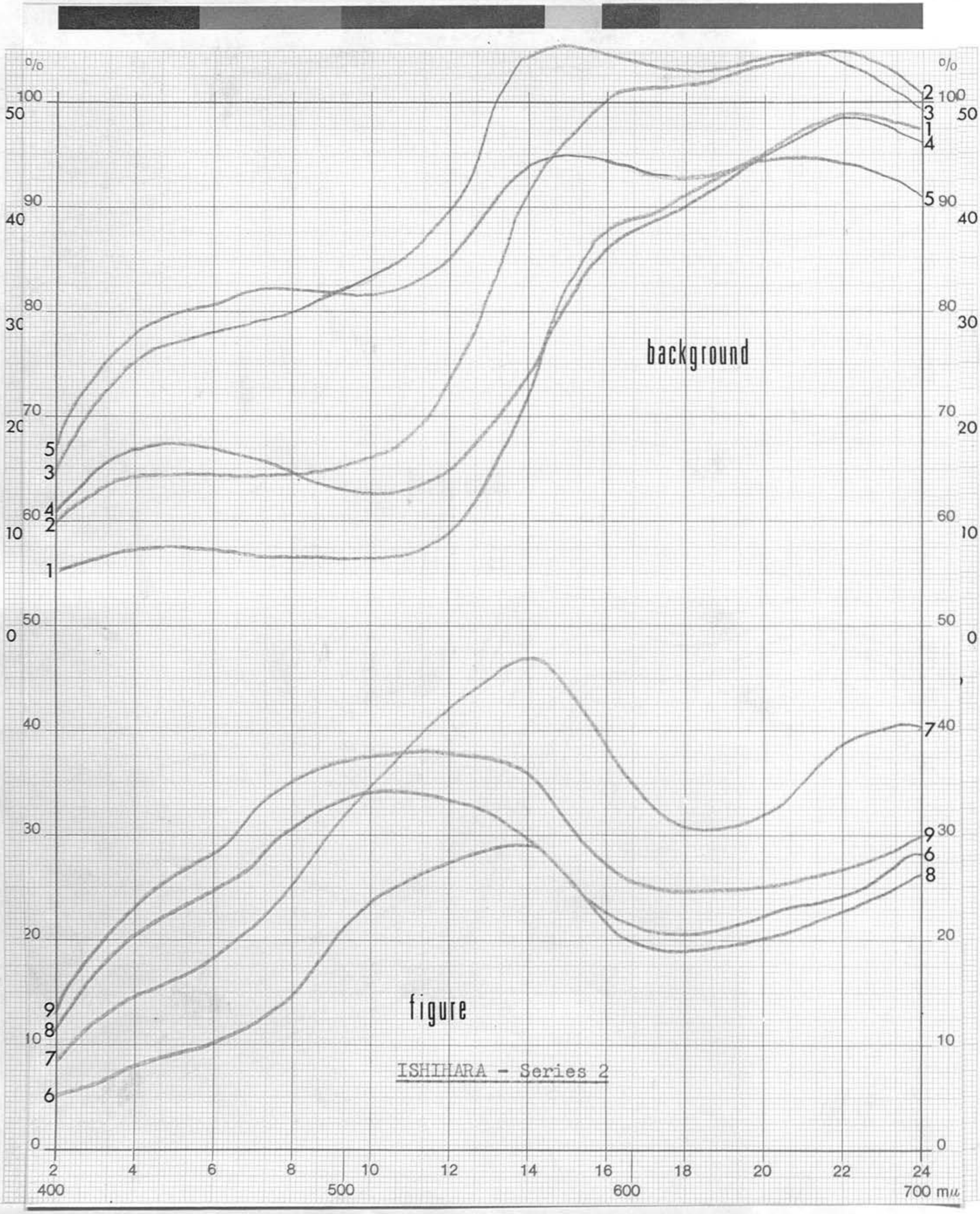
y
x
7.5
43.6
33.8
51.4
8.8
1
2
3
4
5
6
7
8
ISHIHARA
BACKGROUND
FIGURE
+ = ILLUMINANT 'C'
·55
·5
·45
·4
·35
·3
·25
·2
·25
·3
·35
·4
·45
·5
520
510
500
490
480
470
460
450
440
380
580
590
600
610
620
630
640
650
700

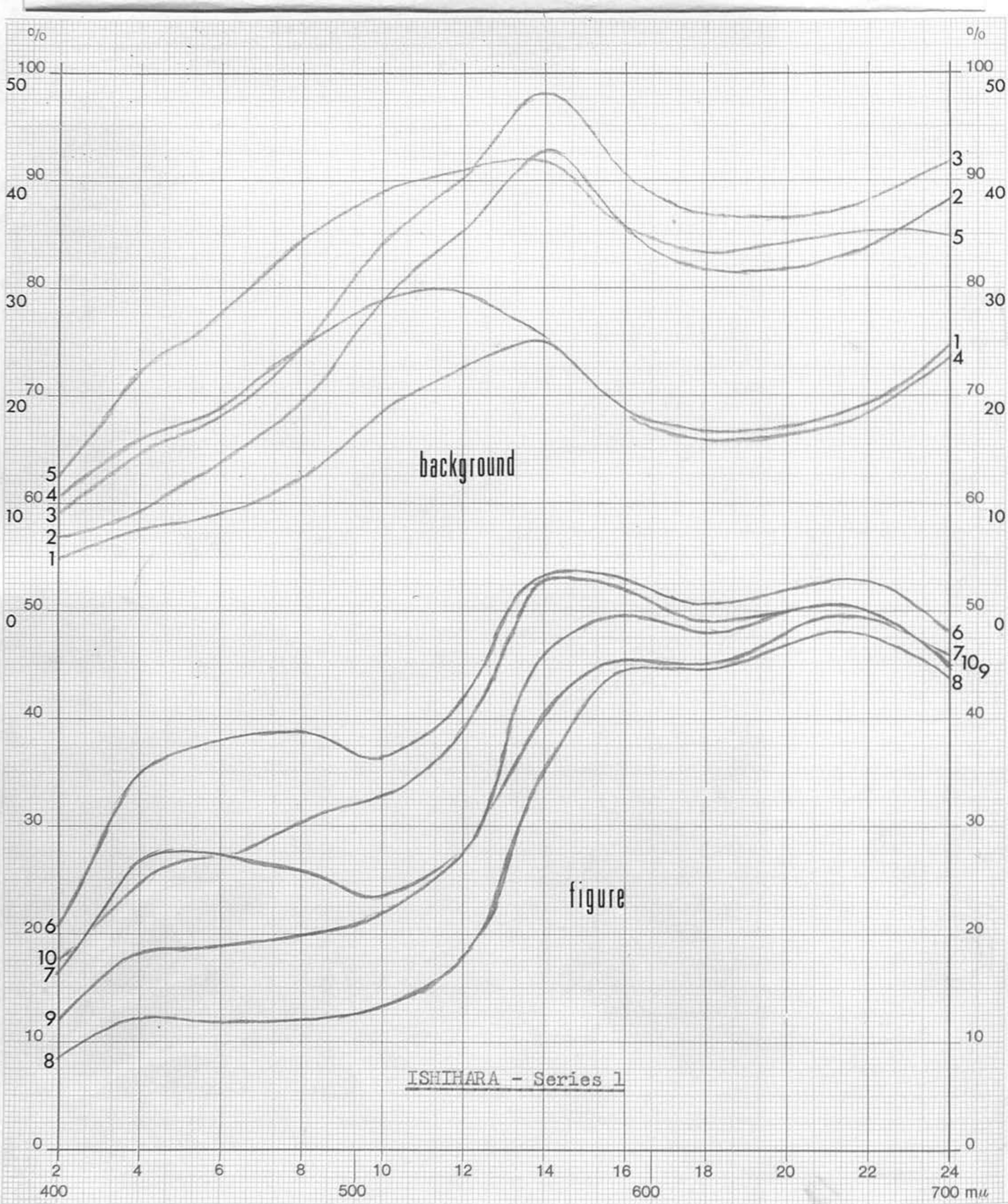
space lie close to the protan and deutan confusion lines. As figure and background colours lie close together in the colour space, fewer colours are needed to make the plates. Whether a subject is a protanope or a deuteranope, he will confuse at least two of the four colours of these plates, and this is enough to prevent him from perceiving a number.

On the other hand, in plates number ten to thirteen of the Ishihara, three colour dots are used for the figure, but the background is composed of two clusters of coloured dots, one on the deutan confusion lines, (in relation to the figure) and the other near the locus for illuminant 'C'. This second series of dots though not placed on the protan confusion lines, must have been designed for that purpose. The misplacing of the element designed to detect the protan types of defect seems to be repeated in the second series using this kind of construction (i.e. plates 14 - 17). In these plates again, the part for detecting deutan defects is well designed, but the part to discover protan types of defect is inadequate.

Perhaps it should be emphasised here that the wrong choice of colours for detecting protan is one of the main features of the ninth edition of the Ishihara and this is even true of the diagnostic plates for detecting outright dichromats.

Another interesting point is that the use of so many colours, more than are theoretically necessary - must be at the basis of the differences between the two tests in the kind of readings that are made. In the Dvorine there is no room left for misreadings other than the expected protan or deutan response while in the Ishihara, because of the extra colours used, there is always the possibility of 'perceiving' a configuration, and thus of misreading a number.


background
figure
ISHIHARA - Series 2
400
500
600
700 mμ



%
100
50
90
40
80
30
70
20
60
10
50
0
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
background
figure
ISHIHARA - Series 1
2
4
6
8
10
12
14
16
18
20
22
24
400
500
600
700 mμ

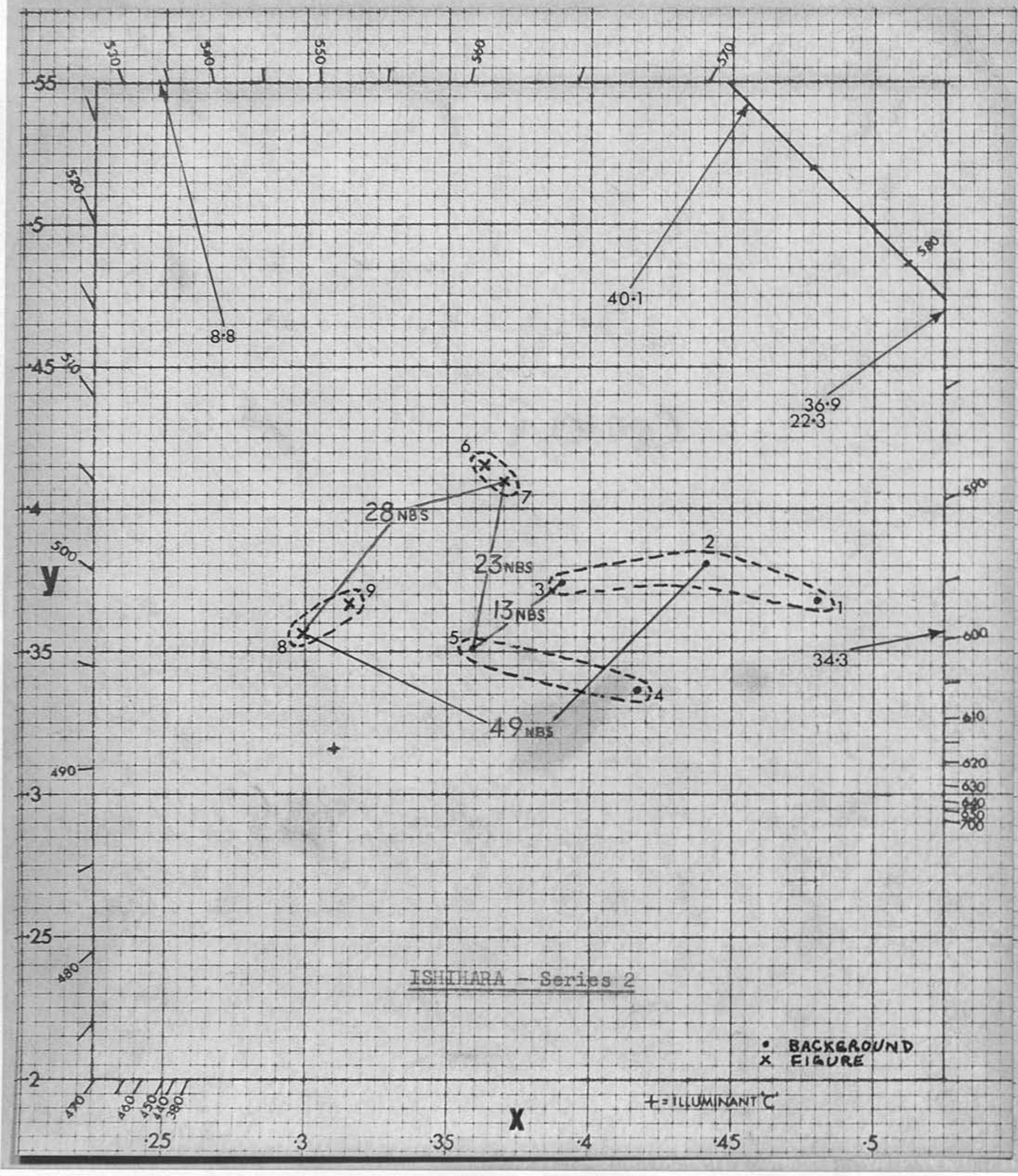
A fuller analysis of the vanishing digit plates will be made when plates ten and eleven of the Dvorine are discussed.

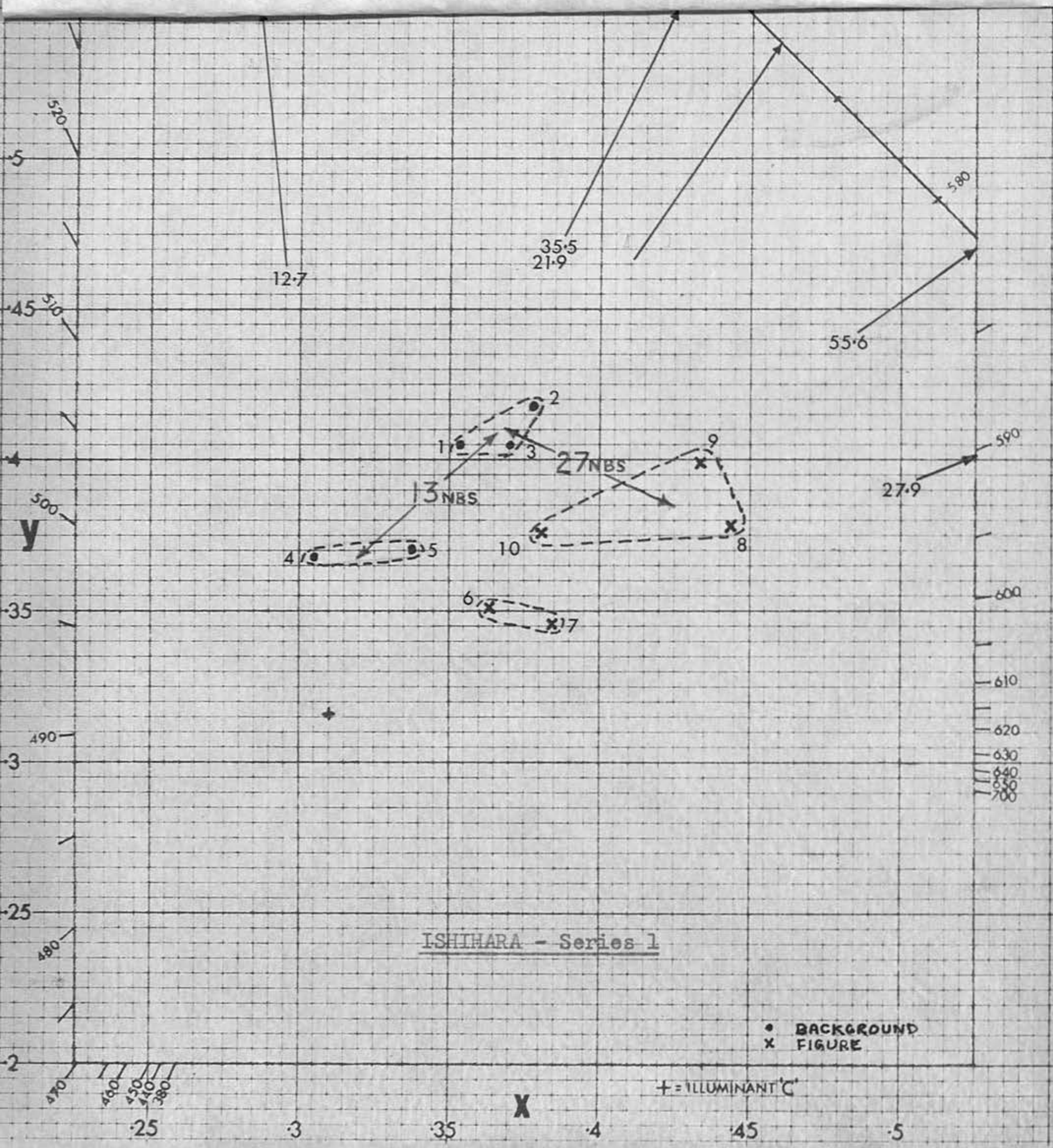
The profusion of colour dots mentioned above affects the first category of plates (that is the transformation type) to an even greater extent than it did the vanishing digit type. These transformation plates appear only in the Ishihara and two series employ this principle - one and two. Both series are similarly constructed. In the first, the figure is composed of coloured dots from the red end of the spectrum and the background of colours from the green end of the spectrum, whereas in the second, the reverse is the case.

Only one of the series is described here, although photometric and colorimetric data are given for both. We are especially interested in Plate 9 reading 74. Nine distinct photometric curves, giving rise to nine distinct sensations of colour can be measured for this plate, four of them are used for the figure, and the remaining five for the background. In the figure, two characteristic types of curve can be distinguished those for colour dots 6 and 7 which have a distinct peak at 560 mμ, and those for dots 8 and 9 which show a plateau around 510 to 520 mμ. In all probability therefore, the four dots were made from only two inks in the printing process, though these varied in weight causing the saturation of the colours to vary and four distinct colours to appear. Again the background is composed of the colours with characteristic reflectance curves; curves 1, 2 and 4 originating from one ink with distinct reflectance peaks at 600 mμ and towards the red end of the spectrum, and curves for colour dots 3 and 5 showing saturation variations only, with rather broader reflectance

GRID No. Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
FIGURE 7	590	27·9	·385	·346	43·7	
8	588		·444	·378	38·3	
6	582		·364	·351	56·8	48·9
9	582	55·6	·434	·399	49·1	
10	579·5		·381	·375	56·6	
BACKGROUND						
2	573		·378	·418	42·4	
3	573		·370	·405	48·8	
1	568	35·5	·353	·405	24·7	41·0
5	568	21·9	·338	·371	39·4	
4	545	12·7	·305	·368	29·5	
Table No. —	COLORIMETRIC DATA FOR ISHIHARA SERIES I					

GRID No. Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
BACKGROUND						
4 9/16:21	599·5	34·3	·417	·337	32·6	
1 9/19:26	594·0		·480	·368	29·1	
2 9/22:19	587·5		·441	·380	45·9	
3 7/23:23	582·0	36·9	·390	·374	59·6	
5 8/17:27	582·0	22·3	·359	·351	61·9	
FIGURE						
7 7/28:29	571·5	40·1	·370	·409	50·0	
6 9/23:35	569·5		·363	·415	31·1	
9 7/27:28	556		·316	·367	49·7	
8 6/18:20	534	8·8	·299	·357	35·3	
Table No. —	COLORIMETRIC DATA FOR ISHIHARA SERIES II					


ISHIHARA - Series 2
• BACKGROUND
x FIGURE
+ = ILLUMINANT 'C'
x
y
8·8
40·1
36·9
22·3
34·3
28NBS
23NBS
13NBS
49NBS



ISHIHARA - Series 1
• BACKGROUND
x FIGURE
+ = ILLUMINANT 'C'
13NBS
27NBS
12·7
35·5
21·9
55·6
27·9
y
x
·25
·3
·35
·4
·45
·5
·2
·25
·3
·35
·4
·45
·5
520
510
500
490
480
470
460
450
440
380
580
590
600
610
620
630
640
650
700

characteristics at the red end of the spectrum where maximum reflectance is around 575 mμ.

The chromaticity diagram showing the loci of the nine distinct colours used in these plates amply illustrate how cleverly these plates were designed, and how advanced they were when they were first brought out. The colours are arranged in four groups; two for the figure and two for the ground. If we go by the dominant wavelength positions for these dots measured from the white reference point (illuminant 'C' in this case) these are seen as distinct colours in terms of hue by the Normal Observer. As a group, the greens have an affinity to each other, and the pink and orange colours also have an affinity to each other. It is this that allows us to perceive, say, a distinctly green figure on a reddish ground.

However, if we consider the case of the protanope or deuteranope who confuses reds and greens, it is evident that his colour affinities will be different. In fact, there is almost a 90° rotation of the colour afinities. Coloured dots 6 and 7 in the figure are now on one of their confusion lines and this can be extended to include colours 32 and 1 of the background while colours 8 and 9 of the figure are now on the other confusion line involved, and this also extends to dots 5 and 4 of the background. As can be seen from the chromaticity diagram on a priori ground, it is the standard observer whom we can expect to read 74, whereas if the coloured dots are arranged according to the confusion lines for either deutan or protan, the number 21 will be read instead.

But, what we may ask, will happen to the subject whose colour discrimination falls somewhere in between these two extremes, whose thresholds for wavelength discrimination are increased over the whole visible spectrum ? Or his difficulties may be confined to certain parts of the spectrum - say between 550 mu to 590 mu - which for him would cause difficulties in perceiving figure and ground clearly and lead to his perceiving a configuration that fits neither the pattern for the normal nor the dichromat.

Plate number nine illustrates such difficulties very well. Besides those who read it as a clear 74 or 21, approximately 50 per cent of apparently normal subjects read it as 71 or as a composite figure where either the first or second digit is misread.

How does this come about ?

A close inspection of the type of misreadings shows that mostly the problem is to read the second digit correctly. If we are to now look at the numeral 4, it can be seen that the stem of this digit is always incorporated in the construction of many of the misread numerals - for example, people may read this plate as 71. The same applies to the stem of the number 7 - which is then incorporated in the misread numerals such as 2 or 9.

From inspection of this plate, it will be seen that the stems in both numeral 7 and 4 are composed of more definite greens - or greeny/blues - while the rest of the figure is composed of yellow-greens.

If the chromaticity diagram is again examined, we find that the stem of the numeral 4 is composed of colours 8 and 9 and if we then refer to table giving colorimetric data for series 11, it will be seen that the dominant wavelength for these two dots is 556 and 534 mu respectively, and they are therefore nearer to the primary green than any of the greens used in the construction of this chart. The same applies to the stem of digit 7. The two colours involved here, namely, dots 6 and 7 have dominant wavelengths of 571.5 and 569.5 mu respectively.

But what about the background colours ? Here again we find two clusters of dots (5 and 3) with a dominant wavelength of 582 mu, while dots 1 and 4 have a dominant wavelength of 599.5 and 594 mu. Although the total distance in terms of wavelength between the extremes of the figure and background is about 65 mu (534 mu to 599.5 mu) the shortest distance is as little as 10 mu, that is between 582 and 571.5 mu.

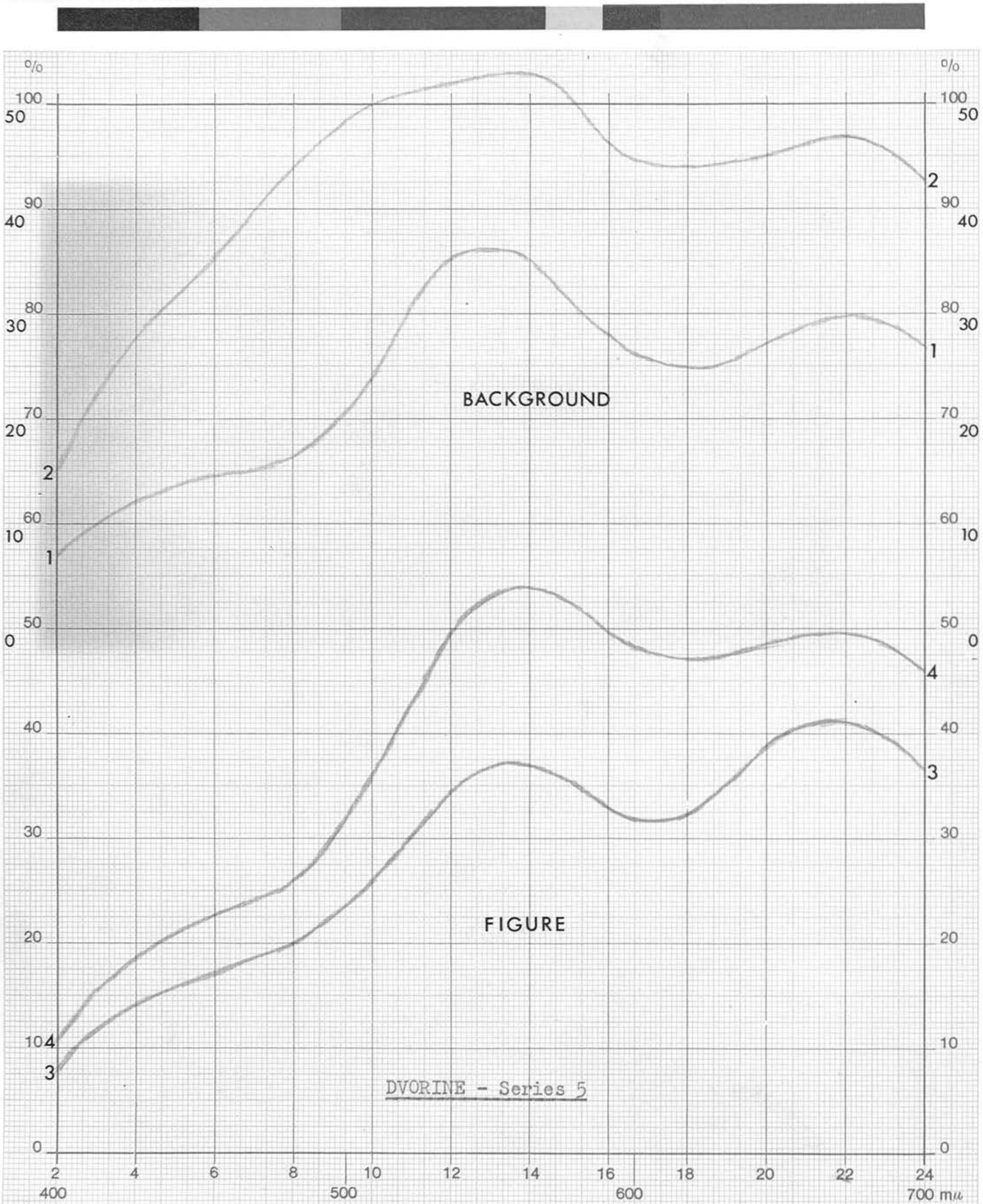
A more accurate measure of colour difference is to express it in terms of N.B.S. units, and here differences in saturation and brightness are also included. Distances are marked in red on the chromaticity diagram. Thus we find that from colour dot No. 8 of the figure to dot No. 2 - the mid point for the background configuration, is about 49 N.B.S. units. Whereas the distance between dot 7 of the figure and 5 of the ground - at 23 N.B.S. units is half the previous distance.

The distance must be even less for the distance between dot No. 7 and dot No. 3, as can be appreciated from the fact that the distance between dot 5 and 3 is only 13 N.B.S. units, and thus, the shortest distance between figure and background for the standard observer is of the order of 13 to 20 N.B.S. units, while greatest differences are around 49 to 55 N.B.S. units. As it happens, the smallest difference between figure to ground in N.B.S. units coincides with the angular part of the numeral 4 and the lower part of the stem in the number 7, and also the surrounding area of this figure, both of the base and of the top of this numeral.

It is safe to assume that in people with minor losses in colour discrimination (due to changes, say in the retinal structure of the ocular media), the colour space in particular areas must shrink and thus would lead to decreased differences between figure and background. This would obviously tend to make the pattern of figure and background less distinct and so the possibility arises of 'inventing' numbers, other than those intended to be read either by the normal or defective person. In the case of plate No. 9, however, reading the first digit as the numeral '2' or the second digit as numeral '1' is also typical of the misreadings found among congenital colour defectives.

The situation found in plate No. 9 where the distance between the figure and the background is as small as 10 N. B. S. units is not common in the Ishihara. Usually the smallest differences in the Ishihara plates are not less than 20 N. B. S. units and are on the average about 30 or 35 N. B. S. units.

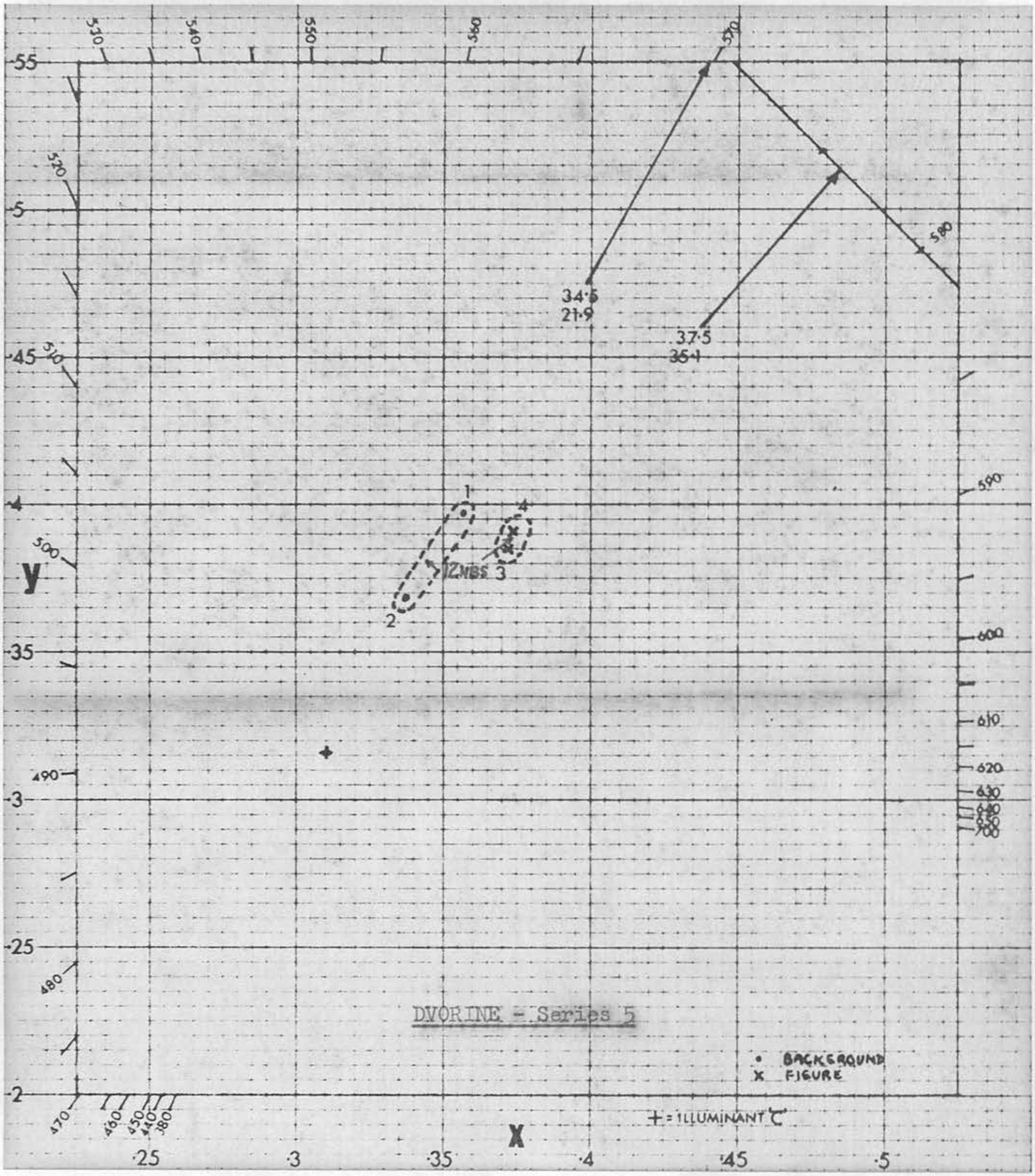
In a previous section dealing with data for the age population tested on the Dvorine and Ishihara, it was found that a greater number of mistakes was made on the Dvorine than on the other test, and it was also stated that the Dvorine showed greater sensitivity to the age variable, judging by the number of misreadings in the various age groups. Now it can be demonstrated that such differences are mainly due to the different structure of these two tests. As has already been mentioned the Dvorine consists mainly of the vanishing type of series and in this connection, it is relevant to point out that on the whole, the distances in Dvorine, between figure and ground, are much smaller than those found in the Ishihara.



BACKGROUND
FIGURE
DVORINE - Series 5
%
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
400
500
600
700 mμ

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Background.							
1	16:20	569·5	34·5	·357	·397	38·8	
2	11:17	569·5	21·9	·338	·369	61·3	
Figure							
3	21:18	576·5	35·[illegible]	·372	·386	42·3	
4.	14:30	576·6	37·5	·374	·391	60·1	

Table No. — COLORIMETRIC DATA FOR DVORINE Series V


DVORINE - Series 5
34·5
21·9
37·5
35·1
12NBS
1
2
3
4
• BACKGROUND
x FIGURE
+ = ILLUMINANT 'C'
x
y
·25
·3
·35
·4
·45
·5
·2
·25
·3
·35
·4
·45
·5
·55
470
480
490
500
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
700
460
450
440
380

With the exception of plate No. 9, most differences in the Ishihara are of the order of 20 or more N. B. S. units.

In the Dvorine, 4 out of the 14 plates have a total distance between the whole figure and the centre of the background of around 12 to 15 N. B. S. units only.

Plates No. 10 and 11 of the series 5 have the least difference of all the Dvorine plates between the colours of the figure and background. This proximity can be observed at a glance, when we look at the photometric data. There are the four curves so typical of the Dvorine, representing the reflectance curves of two pigments each with two different densities, one pigment showing a peak in the yellow part of the spectrum but two levels of total reflectance, and the other pigment used in the construction of the background showing apeak in the yellow/green part of the spectrum.

The chromaticity diagram shows the relative positions of the colours used, and we can see how close the two clusters of colours for the figure and background are to each other. If we now examine the table of colorimetric data for series 5, it will be seen that the background colours have the same dominant wavelength, but show differences in saturation. The same happens with the figure - again there is only one dominant wavelength but slight variations in saturations.

The colour dots also vary in luminosity, both clusters have two 'bright' colours, and two 'darker' colours. This is a simple type of plate and differences in dominant wavelength are only about 6 mu. In terms of colour differences, the

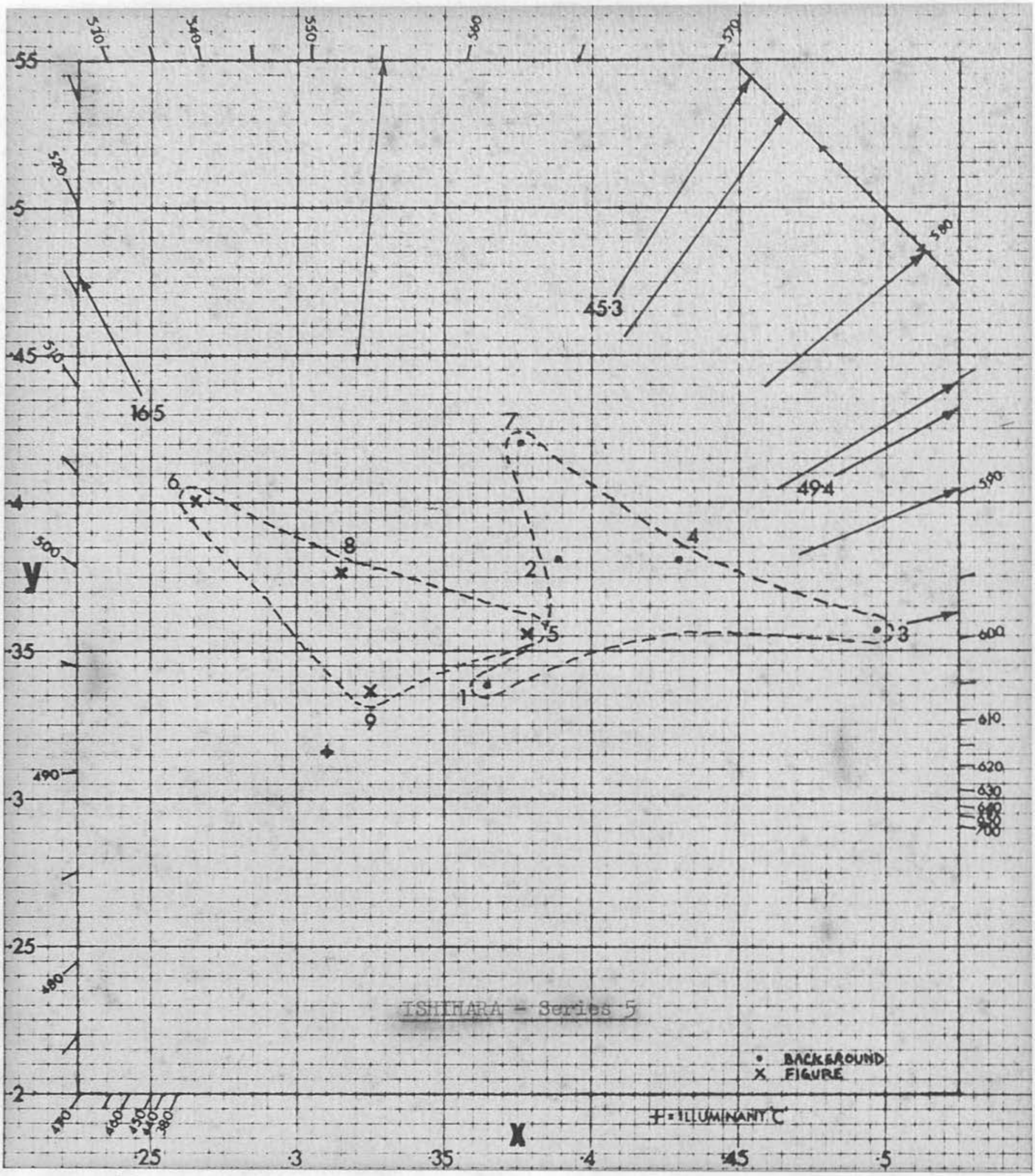
GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Figure							
6	22:22	515.5		.375	.419	37.4	
8	32:23	560.5		.315	.375	42.7	
9	16:27	573.0		.325	.336	35.1	
5	29:29	585.0		.266	.400	21.4	
Background							
7	30:28	572		.379	.356	52.3	
2	31:19	580.3		.388	.381	33.5	
4	15:13	586.0		.430	.381	42.9	
1	18:26	590.0		.365	.338	42.0	
3	19:28	597.5		.497	.357	26.0	

Table No. V COLORIMETRIC DATA FOR ISHIHARA Series V distance between the mid point of dots 2 and 1 of the ground and the mid point of the figure is 12 N.B.S. units. It is much less between dot No. 1 and 4.

Of the two plates (10 and 11) of series 5, plate No. 10 reading 62 has the highest percentage of all misreadings, per plate in the test (almost 40% of those who read it) while plate Number 11 reading 4 has 21%. Both plates have the most sharply defined V or U type of distribution of misreadings for the various age groups. The four plates of the series 4 and 5 account for $4\frac{1}{2}$% of the total misreadings for the total age population.

Before comparing the qualitative plates of the Dvorine and Ishihara test the so called hidden digit plates in series 5 of the Ishihara should first be mentioned.

No photometric curves are presented. There are nine distinct reflectance curves in these plates, and it should be stressed that three of the curves have more or less neutral curves with hardly any peaks, and that the highest of each individual curve is rather lower than these found in the other plates. What is more important, however, is that the colorimetric data points to a unique type of plate construction. Colorimetric data diagram No. V indicates that the spread of chromaticities in these plates is one of the widest. At the one extreme the green is of 515 mu, and at the other, the orangey red is 597.5 mu, this covering a distance of around 82 mu. From this diagram we can see that there is a great variation in saturation, and also that the dots of the figure are less bright than the background in terms of luminosity. The most important element in this type of plate is that the individual colour dots of the figure and background overlap each other in the colour space. It should be remembered that in no other type of plate construction, so far discussed, do we find this overlapping of figure



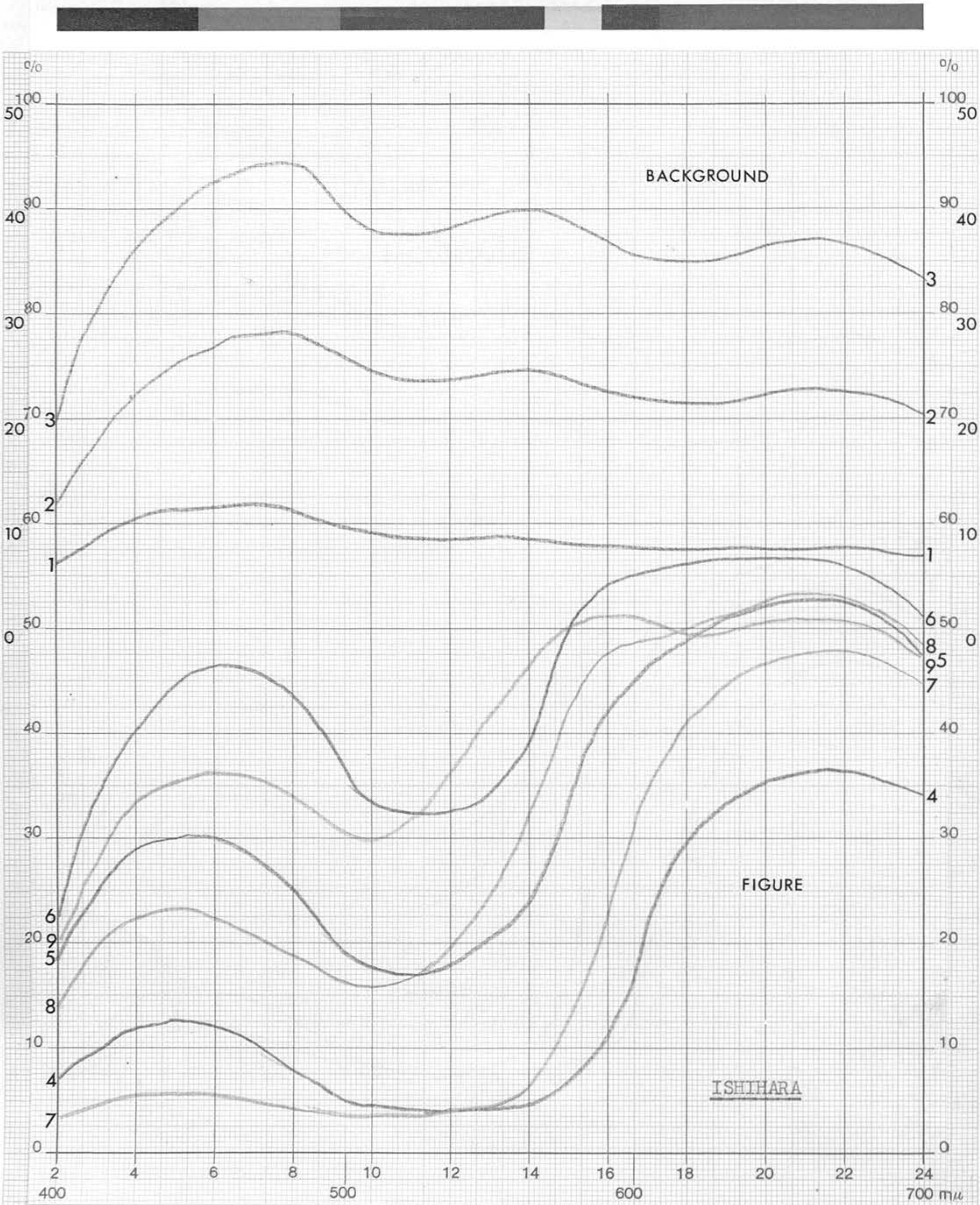
ISHIHARA - Series 5
16·5
45·3
49·4
• BACKGROUND
X FIGURE
+ = ILLUMINANT 'C'
y
x

and background colours. The others all form clusters - and cover colour areas that are quite distinct from each other. Even in series 4 and 5 of the Dvorine this was clearly seen.

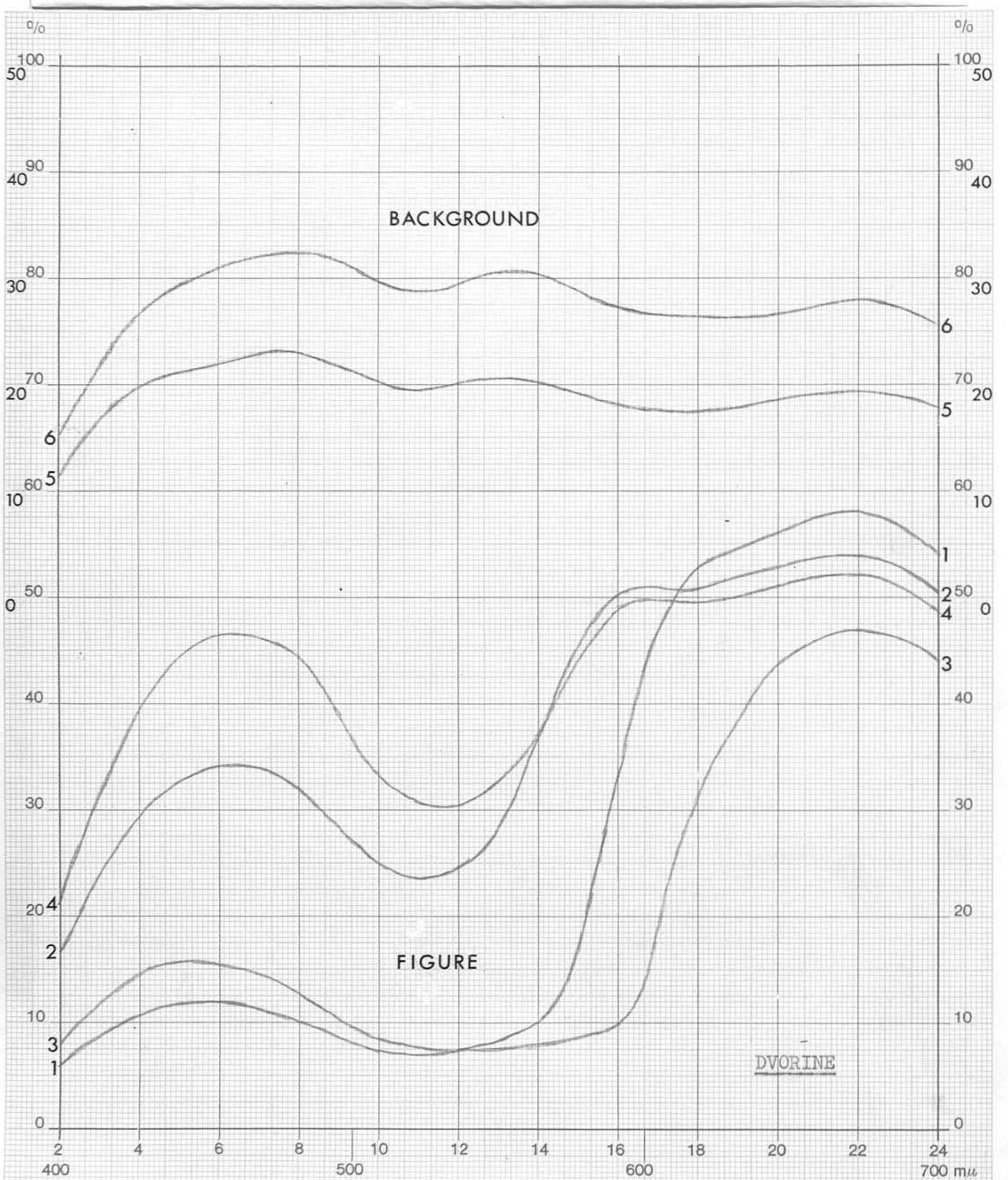
In the 'Hidden Digit' type of plate, however, the colours of dots (9 and 5) of the figure overlaps the dominant wavelength positions for the coloured dots (Nos. 7, 2 and 6) of the background. In addition dot No. 9 of the figure and dot No. 1 of the background have reflectance curves that evoke a neutral sensation in the standard observer. Yet these could appear to be 'colours' depending on the nature of the subject's colour balance, that is to the protanope they might look bluish, to the deuteranope greenish, etc. in colour. Thus to distinguish the colour dots as figure and background, it is necessary to have defective colour vision when almost all the dots of the background will be seen to be lying on one confusion line, and the rest of the colour dots will lie on another confusion line, far enough separated from the first confusion line for it to be possible to distinguish a bluish figure on a rather nondescript background.

The puzzling thing, of course, is that young subjects with normal colour discrimination in the 18 - 25 age group, may also read a number in this series.

Perhaps this could be explained by the fact that the 'balance' between the red and blue ends of the spectrum is quite different for normal subjects of this age group than for any other age group. Warburton's study showed that subjects in the 20 - 30 age group gave a higher match for illuminant 'B' in terms of colour temperature than any other age group. (Later, evidence will be shown from the anomaloscope that the most modal ratios for the three equations are different for the 16 - 25 age group). If this were true, it would then be possible to explain



BACKGROUND
FIGURE
ISHIHARA
%
400
500
600
700 mμ



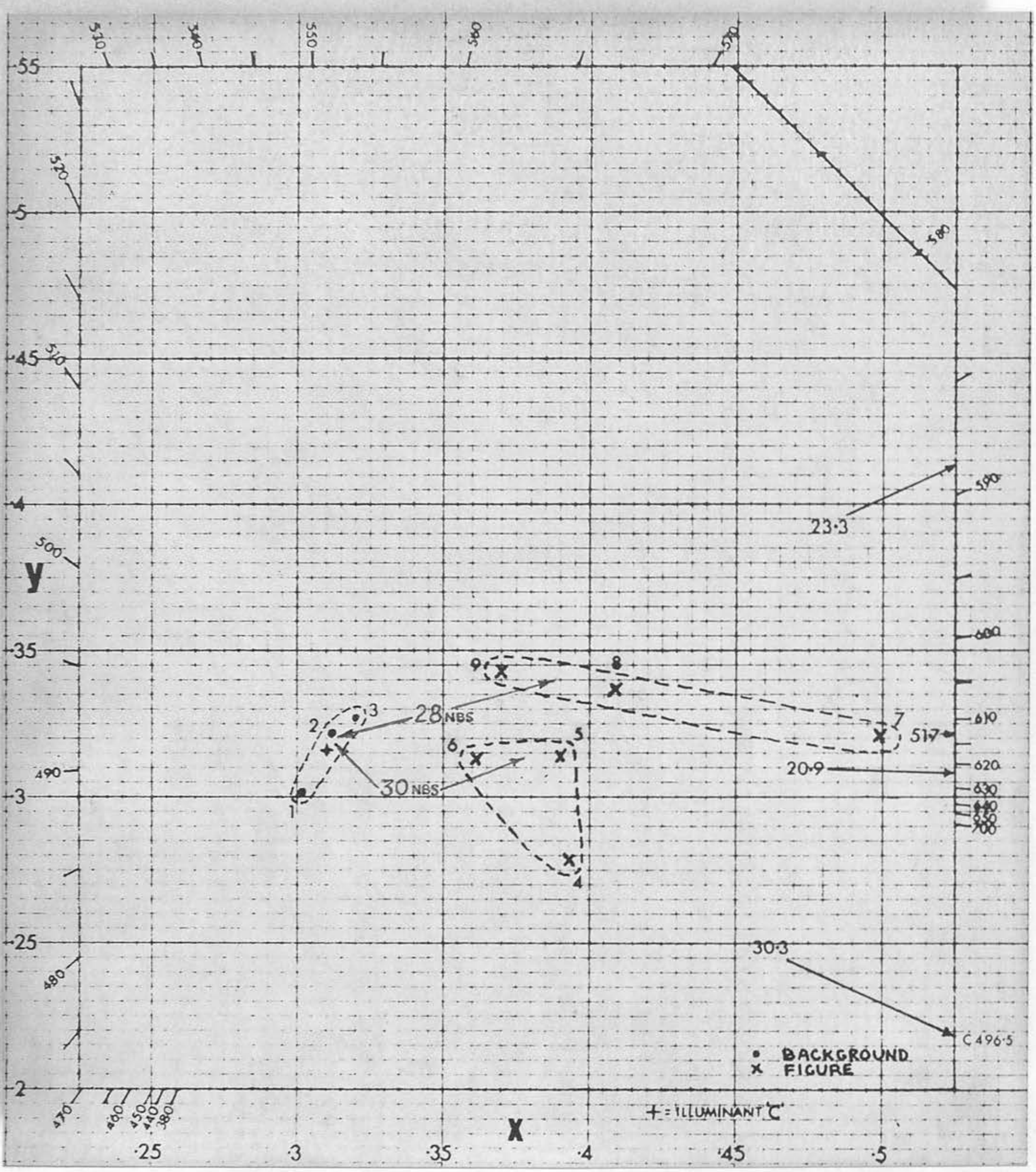
BACKGROUND
FIGURE
DVORINE
%
100
90
80
70
60
50
40
30
20
10
0
2
4
6
8
10
12
14
16
18
20
22
24
400
500
600
700 mμ

why young subjects can see the hidden digits, for the separation of the coloured dots in such plates would become much greater than they were for the standard observer. The greens at 515 mu would become more bluish, the colour dots around illuminant 'C' would be nearer to that illuminant (especially those with neutral reflectance curves) and on the other hand, the orange and pinkish dots of the background would become more yellowy-orange. It is possible to test this hypothesis by means of a visual colorimeter if the individual dots were to be measured by subjects of that particular age group, and then by others in the 30 to 40 age group.

Now we come to the last type of plate used in the tests. This is the qualitatively diagnostic type used for detection of dichromats. Both Ishihara and Dvorine employ such plates, there being four of them in the Ishihara, and two in the Dvorine. Basically, this type of plate is constructed on the principles of the vanishing plates, except that differentiation is built in so that only part of the composite figure is unreadable to a given dichromat. This principle is employed in both tests, and essentially it is two vanishing plates rolled into one.

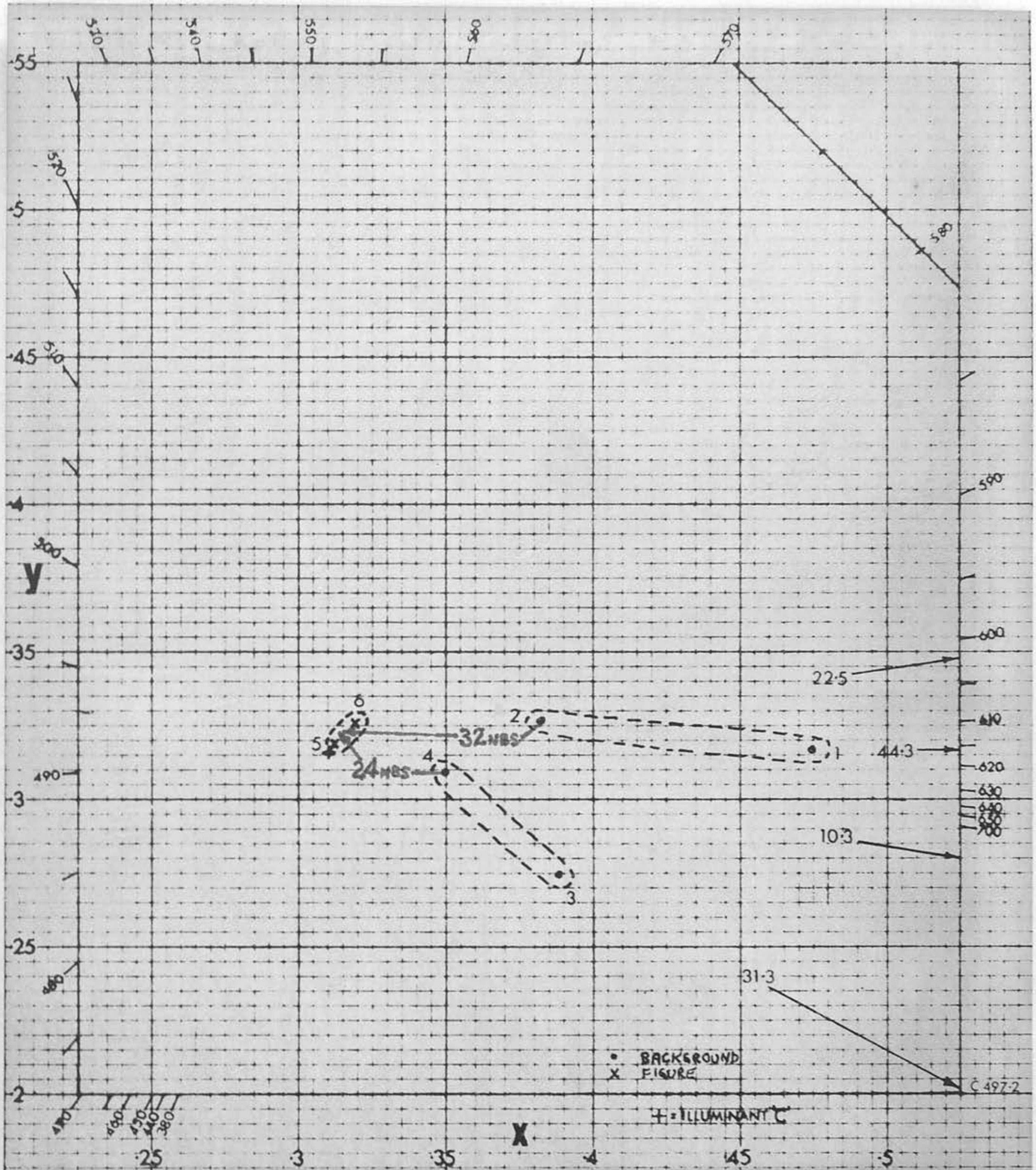
From the diagram of the photometric curves for both tests we see that, as before, there are more colours used in the Ishihara than in the Dvorine plates. In the Dvorine, there are four curves for the composite numerals of the figure, and two reflectance curves giving the characteristics of the background. On the other hand, the Ishihara has three curves for the background, and six for the figure.

Spectrophotometrically, the curves for the background seem to be very similar in both tests, except for colour dot No. 3 of the Ishihara, which is not so



y
X
28 NBS
30 NBS
23·3
51·7
20·9
30·3
C 496·5
• BACKGROUND
x FIGURE
+ = ILLUMINANT 'C'

ISHIHARA



BACKGROUND
X FIGURE
= ILLUMINANT C
32 NBS
24 NBS
22·5
44·3
10·3
31·3
C 497·2
x
y

DVORINE

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Background							
1	18:16	'c'	–	·301	·301	12·1	
2	21:29	'c'	–	·312	·322	30·7	
3	19:28	'c'	–	·320	·327	47·3	
Figure							
PROTAN							
9	23:19	589	23·3	·370	·343	51·8	
8	17:21	598		·409	·337	39·4	
7	37:36	612	51·7	·499	·321	19·4	
DEUTAN							
5	18:26	622	20·9	·390	·314	37·4	
6	29:27	622		·361	·313	52·4	
4	27:34	c 496·5	30·3	·393	·279	22·2	
Table No. _		COLORIMETRIC DATA FOR ISHIHARA Series VI					

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Background.							
5	12:27	'c'	–	·312	·319	25·8	
6	11:17	'c'	–	·319	·326	37·3	
Figure							
PROTAN.							
1	8:17	616	44·3	·475	·317	28·0	
2	11:24	602	22·5	·383	·327	48·3	
DEUTAN							
3	9:26	c 497·2	31·3	·389	·275	22·6	
4	10:29	c 493·0	10·3	·350	·309	52·2	
Table No.		COLORIMETRIC DATA FOR DVORINE Series III					

neutral. The pigments used for making the deutan figure, exhibit characteristic curves of a minus green type, the greatest absorption being at 535 mu. There is also a similarity between the two tests in the construction of the figure for protans. However, it is surprising to see that in both tests, pigments in the characteristics close to those of the deutan figure were used for the protan figure (see Nos. 2 and 4 of the Dvorine; 6 and 9, and 8 and 5 of the Ishihara).

Referring now to the colorimetric data sheet for series 6 of the Ishihara and series 3 of the Dvorine, we note that the background dots for both tests are very close to the position of illuminant 'C'. For the protan figure, the Dvorine uses colour dots with dominant wavelengths of 616 to 602, while in the Ishihara two orangey dots are employed. For the deutan figure, both colour dots used in the Dvorine are in the purple region, while the Ishihara has only one colour dot with a complementary wavelength.

Examination of the chromaticity diagrams proves rewarding. Here we see in both the Ishihara and Dvorine the colour dots for the deutan figure in relation to the background lie on the deutan confusion lines, but in contrast to this, the colours of the protan figure are quite distinct from the background configuration of the Ishihara, and only just touches it for the Dvorine. Because some of the deutan colour dots lie on the protan confusion lines, <u>the figure the protanope is supposed to see</u>, becomes indistinguishable to him.

From this analysis, differences in the performance of deutans and protans on these plates would be expected. The deutan would be unable to read the deutan figure, but should always be able to read the protan figure, especially in

ISHIHARA

PROTAN

No.	AGE	−6	−2	−5	−6
1	16	−	42	−5	−
2	20	7	4	55	0−
3	19	18	42	35	35
4	20	−	−	−	−
5	10	−	−	−	−
6	33	−	−	5	−8
7	36	−	−	3	−
8	20	−	−	−	−
9	14	−	45	−	−
10	18	26	−	35	26
11	16	−8	−2	−6	−
12	10	−6	−2	−	−6
13	8	−8	−2	−5	−8
14	10	−6	−2	−5	−6
15	60+	−	−	−	−
16	60+	−	−	−	−
17	60+	−	−	−	−
18	60+	−	−	−	−

DEUTAN

No.	AGE	2−	4−	3−	9−
1	20	2−	4−	3−	9−
2	27	2−	4−	3−	9−
3	16	2−	4−	3−	9−
4	15	2−	4−	−	9−
5	15	−	5−	3−	−
6	24	2−	4−	3−	9−
7	23	2−	4−	3−	9−
8	10	2−	4−	−	9−
9	23	2−	4−	3−	9−
10	−	2−	4−	3−	9−
11	41	2−	4−	3−	9−
12	50	2−	4−	3−	9−
13	60+	2−	4−	3−	9−
14					
15					
16					
17					
18					

TABLE No. ___

DVORINE

PROTAN

No.	AGE	−5	−6
1	16	−	−
2	20	−	−
3	19	−	−
4	20	−	−
5	10	−	−6
6	8	−	−
7	10	−5	−6
8	10	−	−
9	33	−5	−6
10	36	−	−
11	14	−	−
12	16	−	−
13	15	−5	−
14	14	−5	−6
15	16	−5	−
16	60+	−	−
17	60+	−	−
18			
19			
20			

DEUTAN

No.	AGE	9−	2−
1	16	9−	2−
2	19	9−	2−
3	23	9−	2−
4	19	9−	2−
5	20	9−	2−
6	22	9−	8−
7	31	−	−
8	41	2−	−
9	50	−	−
10	23	9−	2−
11	15	9−	2−
12	18	9−	2−
13	8	9−	2−
14	40	−	2−
15	16	−	2−
16	15	9−	2−
17	16	−	−
18	15	−	−
19	60+	−	−
20	60+	−	−

TABLE No. ___

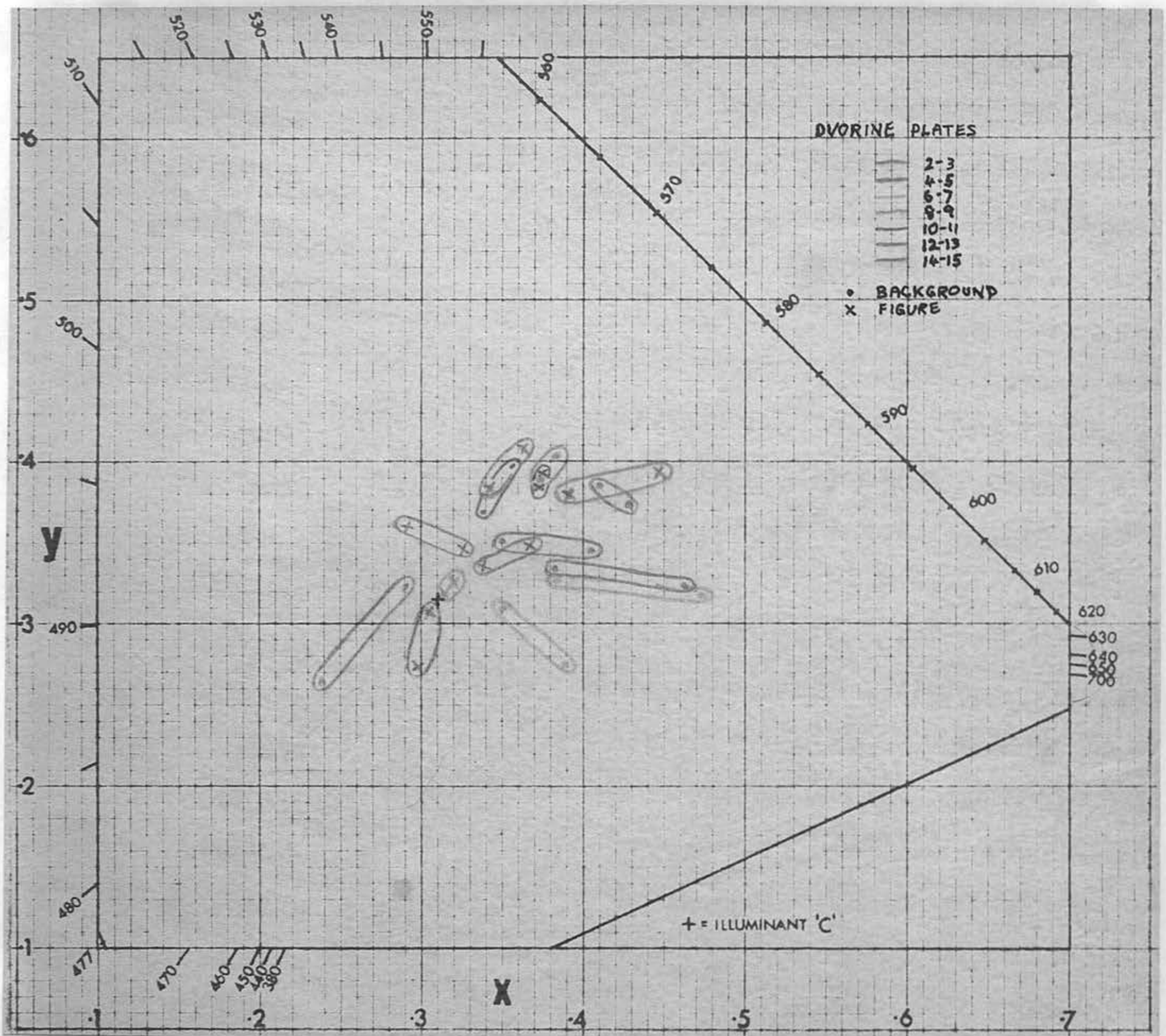
the Ishihara. Because of the proximity of some of the deutan colours to the colour area of the protan figure, we would expect that protans would find it difficult to read even the deutan figure. To the protan, the elements in the background will stand out distinctly as the dominant configuration, thus leading to failure to perceive a figure.

In terms of confusion lines, the deutan chart of the Ishihara is ideal, while the Dvorine is less so. This can be confirmed by actual readings for about 15 deuteranopes and 15 protanopes.

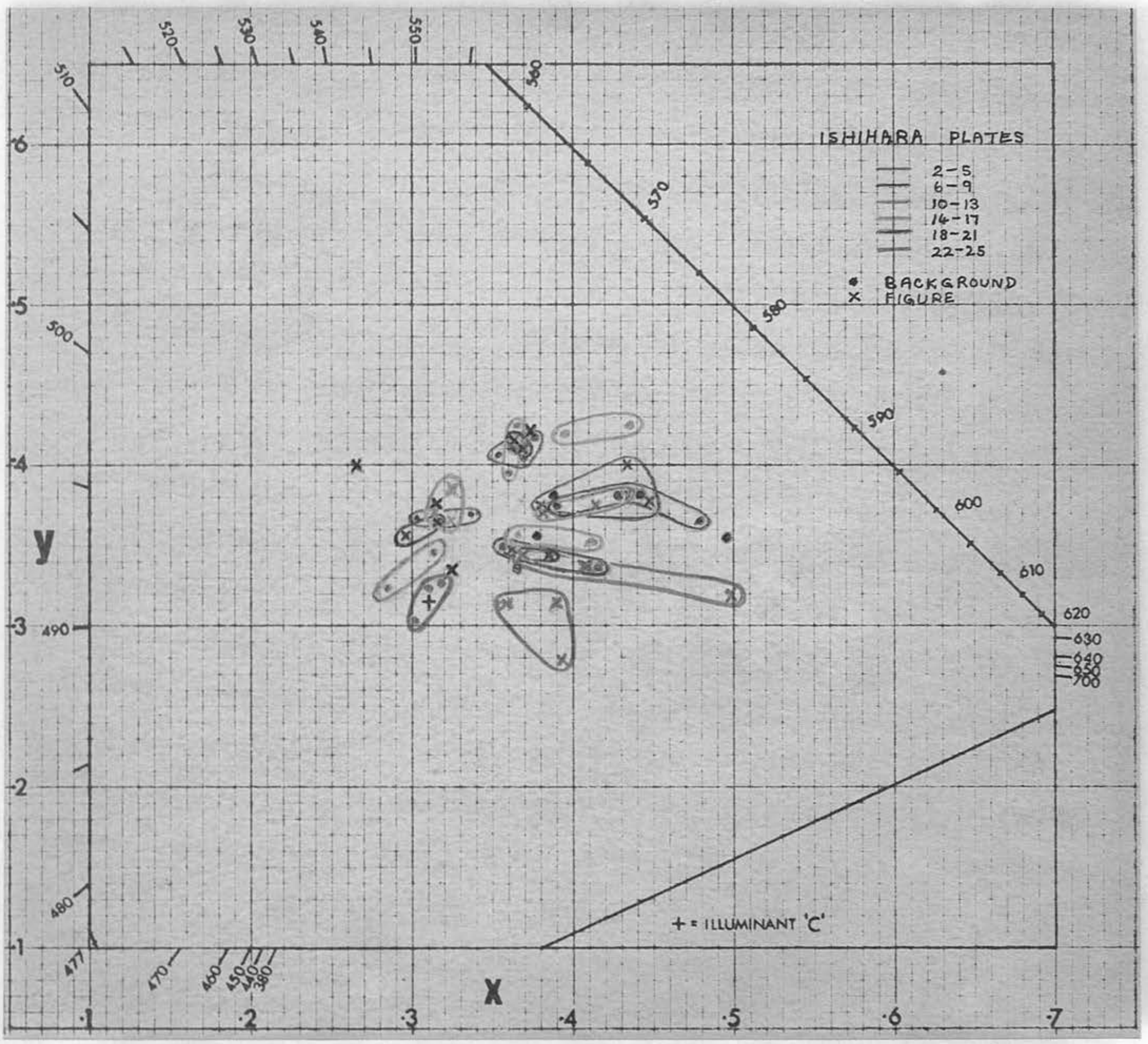
The results refer only to the diagnostic plates of the Ishihara and Dvorine, (facing this page). Note that for the Ishihara diagnostic plates, only one of the 18 protanopes responded in the expected manner as stated in the manual. If all gave the expected answers, 72 of the deutan figures would have been read, whereas only 16 were read 'correctly', that is approximately a quarter of the expected total.

Let us now look at the performance of the thirteen deutan subjects on the same plates. Ten of them gave all 'correct' responses for the four plates. In terms of total responses, 47 of the possible 52 were realised, that is, over 90 per cent of the expected readings.

A similar picture emerges when the Dvorine is examined. Results are shown for 17 protanopes - the correct response is 5 and 6 and only three of this group gave this type of reading. Of the possible total of 34, nine numbers were read correctly, only a little more than a quarter of the expected response. There were twenty one deutans, and of these ten read 9 and 2 correctly. In terms of total readings, 25 out of the possible 42 were



DVORINE PLATES
2-3
4-5
6-7
8-9
10-11
12-13
14-15
• BACKGROUND
x FIGURE
+ = ILLUMINANT 'C'
y
x

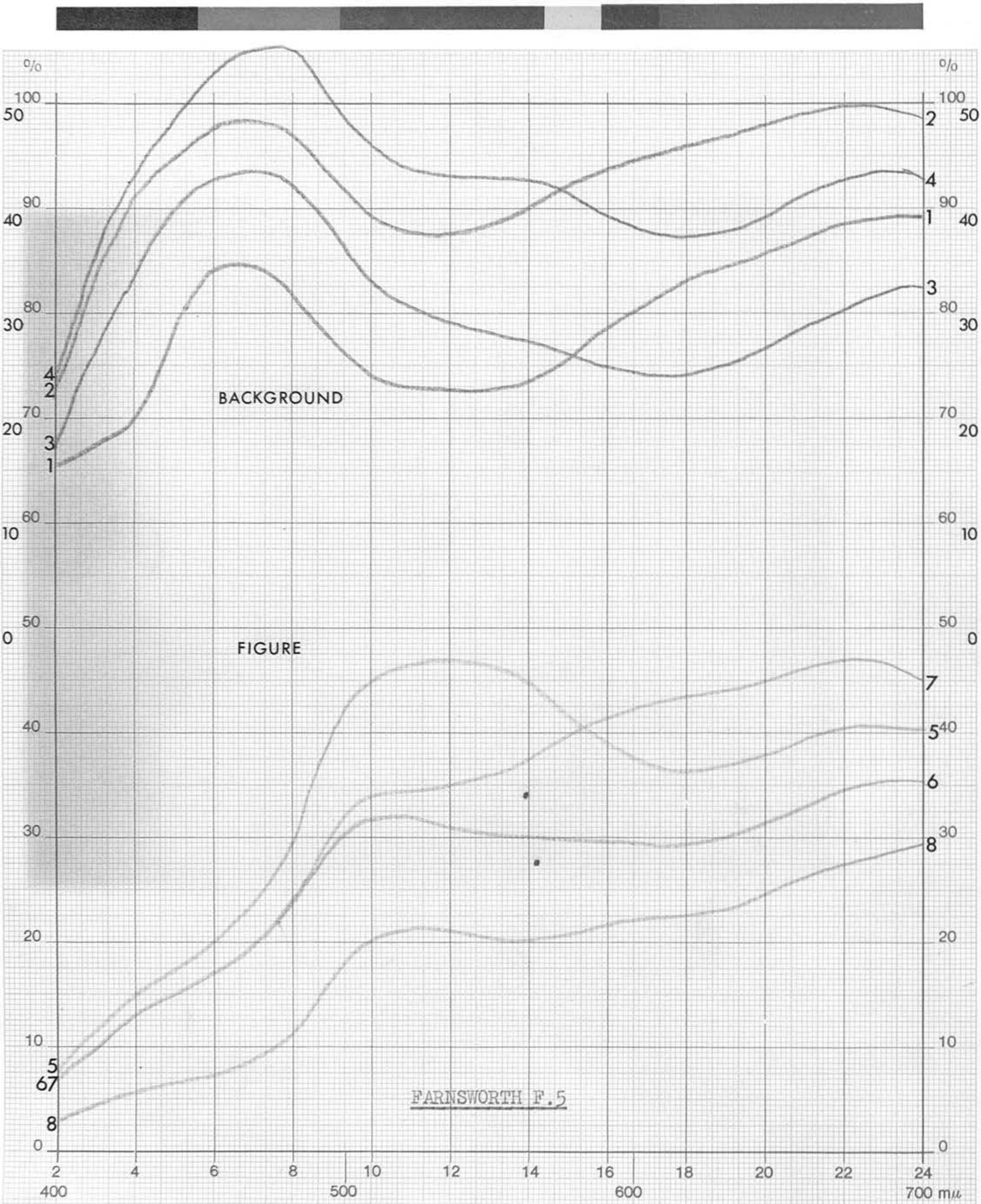

ISHIHARA PLATES
2–5
6–9
10–13
14–17
18–21
22–25
• BACKGROUND
× FIGURE
+ = ILLUMINANT 'C'
y
x
·1
·2
·3
·4
·5
·6
·7
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
700
500
490
480
477
470
460
450
440
380

read, representing a percentage of 60 for 'correct' readings.

It should be pointed out that these diagnostic plates would be tremendously improved if more purple coloured dots were chosen for the deutan figure. If in terms of the colour space, the two areas involved in the protan and deutan tests were further apart, the efficacy of both the Ishihara and Dvorine qualitative plates would be vastly increased.

Finally, there are two chromaticity diagrams, one for the Dvorine and one for the Ishihara, showing the position in the colour space for all the plates of a given test. Note the positions of figure and background in plate 14 and 15 of the Dvorine. These plates have been regarded as a measure of blue-green discrimination, by some research workers, but as can be seen, the position of the dots does not substantiate this inference. Little correlation was found in this research between performance on this plate and losses of colour discrimination in the yellow-blue and violet-blue/green equations of the anomaloscope. As a matter of fact, judging by the position of the 'figure and background' these two plates could be regarded as excellent examples of the vanishing type, designed to detect major red-green defects, since the distribution of the individual colour dots would fit either protan or deutan confusion lines.

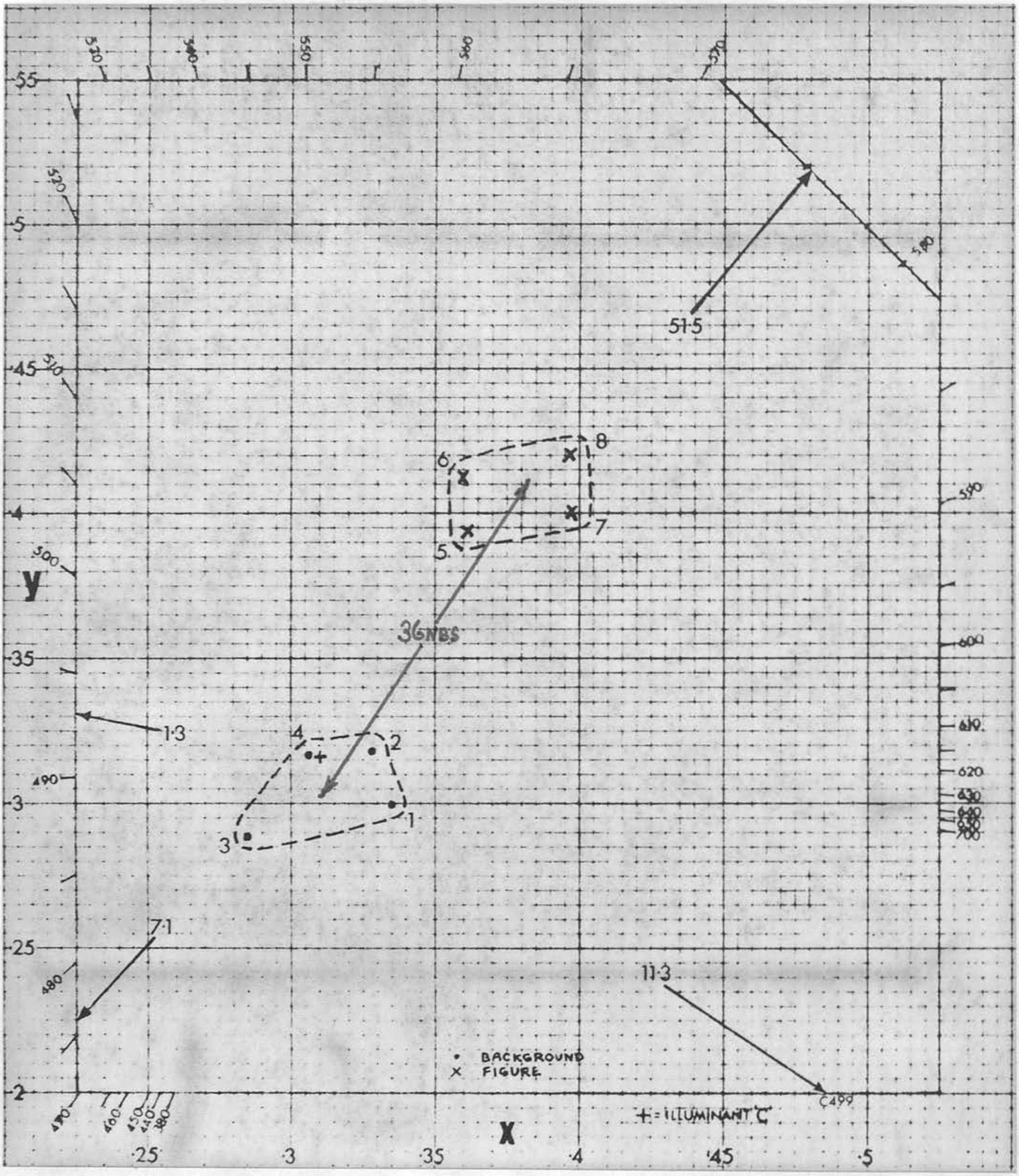
III. Tritan Plates - Of the two Farnsworth plates, F. 5 and F. 2, we noticed previously that the F. 5 plate gave very clear cut results - a decided U type of curve for the incidence of those who did not recognise it at all, and an inverted V type curve for those who saw it distinctly as the figure 5. So both young and older subjects clearly found it difficult to read.



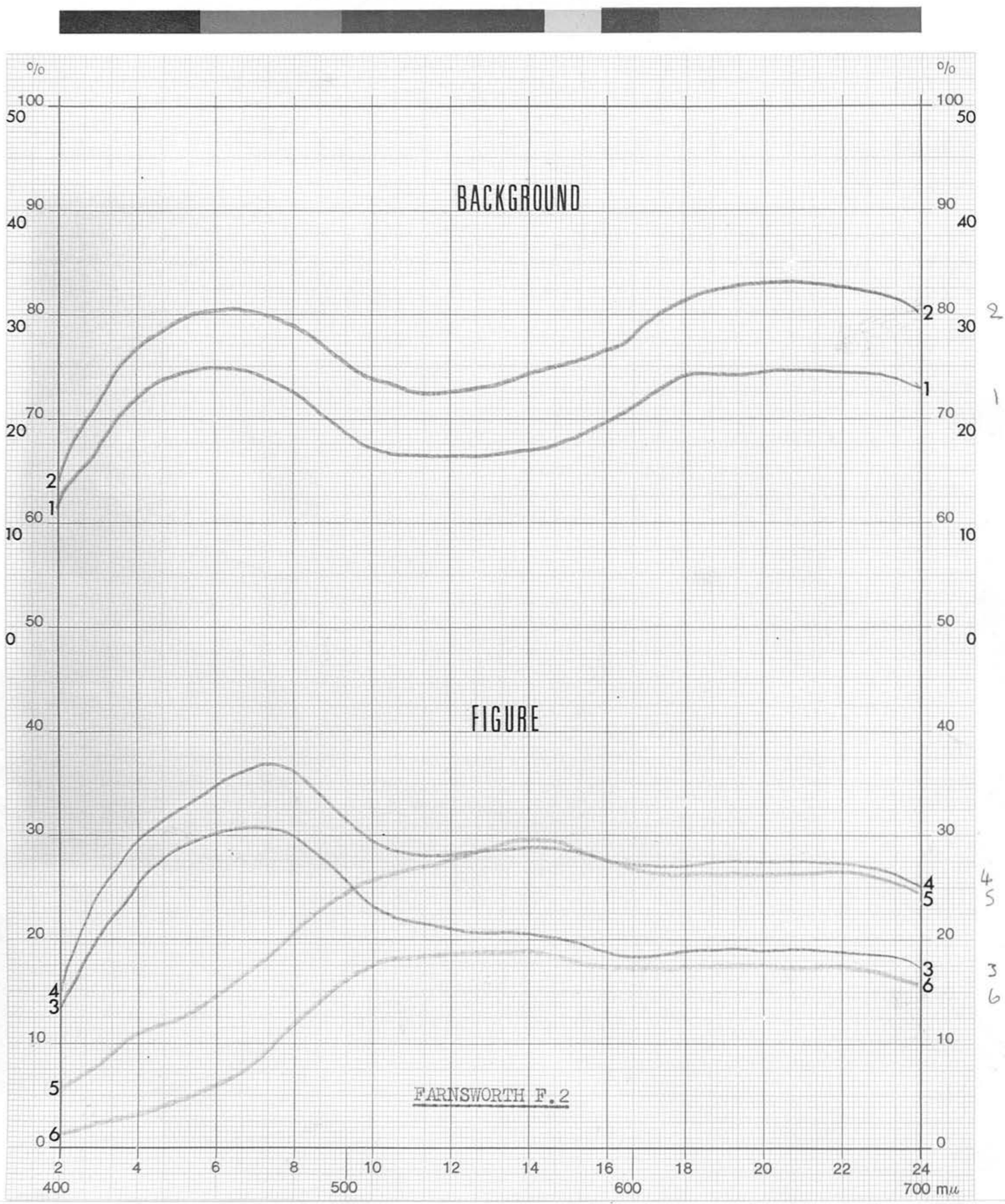
BACKGROUND
FIGURE
FARNSWORTH F.5
%
400
500
600
700 mμ

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Background							
1	17:18	'C'	11·3	·3352	·2995	32·5	
2	14:15	'C'		·3283	·3176	50·2	
3	14:21	'C'	7·1	·2850	·2885	35·0	
4	16:24	'C'	1·3	·3061	·3166	52·8	
Figure.							
5	14:28	573·0		·3608	·3941		
6	18:27	568·0		·3604	·4120	52·9	
7	10:27	578·0		·3971	·3999		
8	10:14	575·3	51·5	·3966	·4197	27·8	

Table No. _ COLORIMETRIC DATA FOR Farnsworth F. 5.



FARNSWORTH F.5



BACKGROUND
FIGURE
FARNSWORTH F.2
%
700 mμ
400
500
600
2
1
4
5
3
6

changes in 'W' are towards the lower colour temperature, it is reasonable to assume that in terms of total colour difference, the distance between the figure and background will change as the subject's 'W' moves towards illuminant 'A'.

The second point is also relevant since from studies of acquired dyschromatopsias it was found that losses of discrimination around 575 mu are quite frequent, and that these lead to the emergence of a neutral point there. Thus, it should not surprise us that so many subjects over 50 see no figure. To them the plate must appear as a mass of yellowish-olive or brownish dots, and the blues will not be seen (except perhaps for the colour of dot No. 3). Note also that the confusion loci for copunctal point of 460 mu are closer to the relative positions of the figure and background than are the confusion loci for tritans.

Farnsworth F. 2 plate presents a different picture. Only in the last two age groups does the number of those who see the two squares fall as low as 50 per cent. Though there are slight variations over the other age groups, these are not significant at all, since around 90 per cent of the subjects can recognise both squares, and see the green as the dominant one.

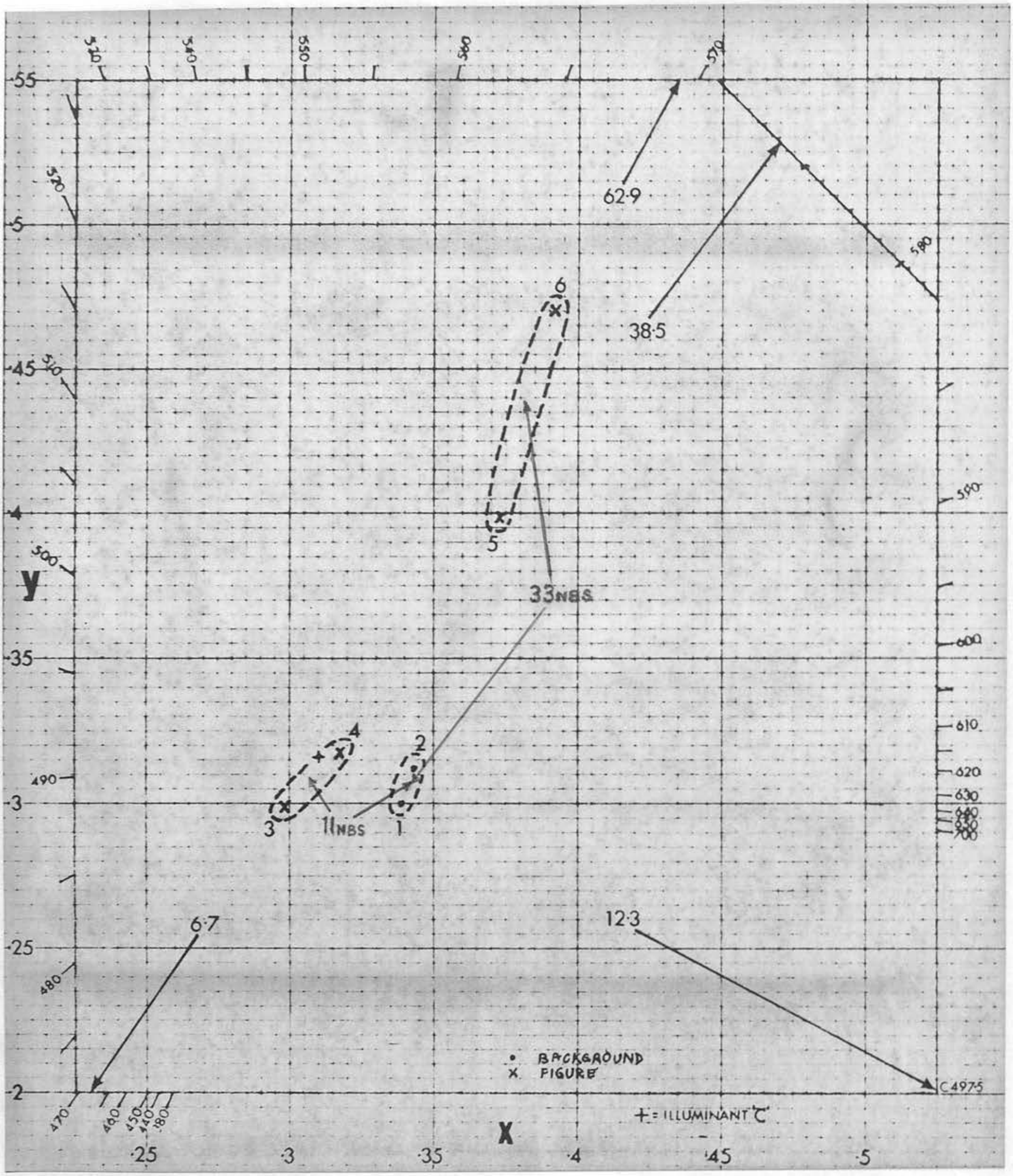
Let us see if the colorimetric analysis indicates why this is so.

Photometrically, there are 6 distinct curves - two for the background, and four for the figure. The background must be made from one pigment with two different densities. Curves for the figure show that only two inks were used, one bluish with two weights, and one yellow-green again with two different weights. The overall reflectance levels for all curves is around 30 per cent.

Referring now to chromaticity diagram giving colorimetric data

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Background.							
1	23:18	C 497.5	12.3	·3385	·300	24·3	
2	26:18						
Figure							
Blue 3	18:16	468	6·7	·2985	·2989	27·2	
4	21:16	C		·3167	·3174	36·5	
Green 5	23:21	574	38·5	·3727	·3986	34·5	
6	26·21	569	62·9	·3915	·4703	22·5	

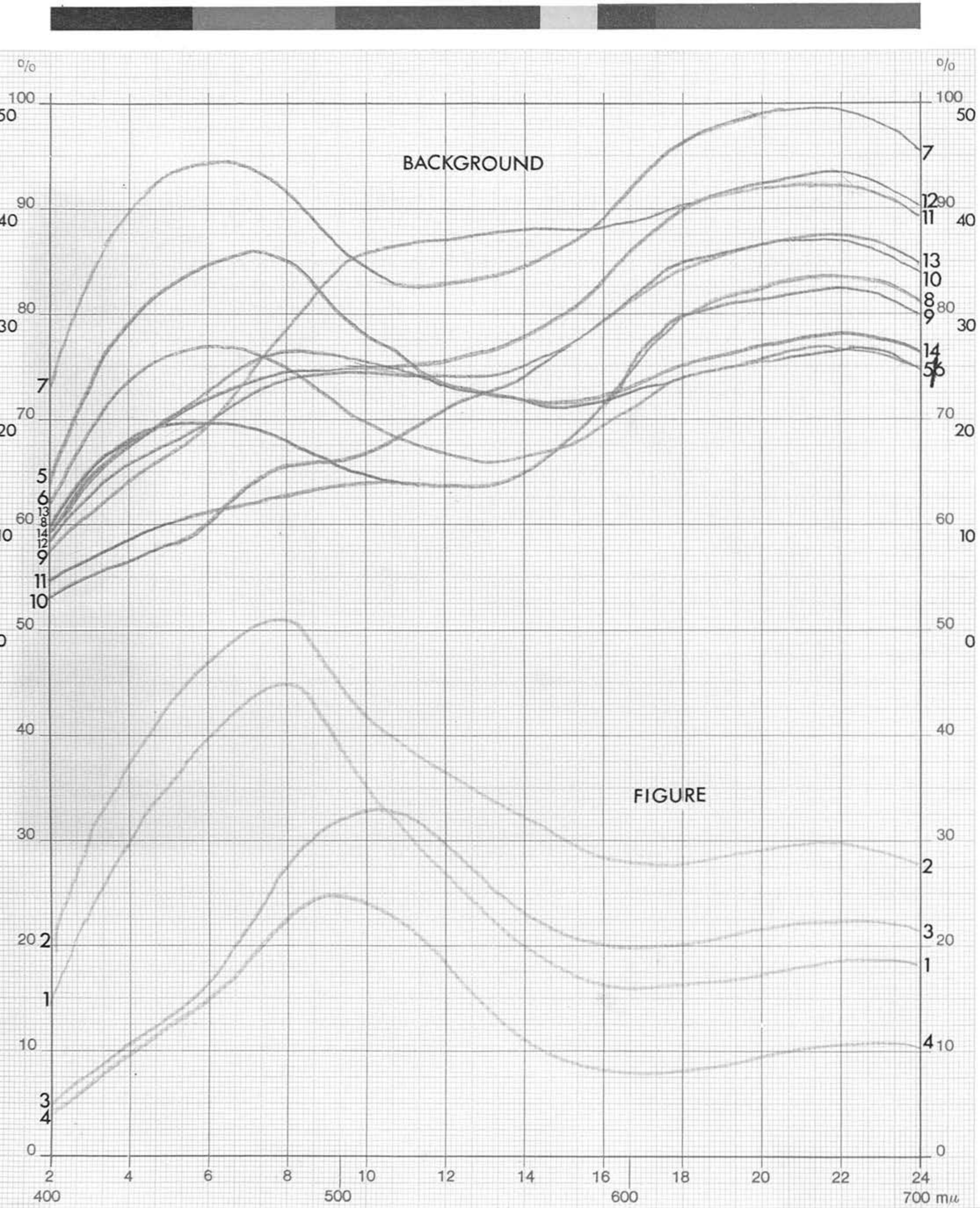
Table No. __ COLORIMETRIC DATA FOR Farnsworth F. 2



FARNSWORTH F.2

(Table No. F.2) the first striking factor is that the general outlay of the clusters of colours resembles one of the qualitative tests for red-green defects, if the position of the cluster for the background is compared with the bluish dots of the figure. The coloured dots for the green square are quite distinct from the two other areas, and it could reasonably be expected that if the green square is ignored, the position of the blue square in relation to the purple background would make an ideal red-green test, and this has proved to be the case, for as we have already stated in the previous section, this plate is used for confirming the existence of major red-green defects. It is a curious fact, that no subject with deuteranomalous or protanomalous vision fails to see the blue and green squares, but all extreme deuteranomolous and extreme protanomalous and all deuteranopes and protanopes are unable to see the blue square but have no difficulty in seeing the green square, though they may misname its colour.

In designing this plate, the intention must surely have been to detect Tritan types of subjects, who would be expected to confuse dots 1 and 2 of the background with dots 5 and 6 of the green figure, leaving dots 3 and 4 to be seen as a distinct blue square. Referring back to the diagram, let us apply the transparency giving the outline of the tritan confusion lines. It will be seen that the placing in the colour space of the background colours goes too far into the purple area of the chromaticity diagram. One wonders then what this plate really can test ? Perhaps it would be safer to regard it as a test of red-green discrimination. In comparison with this, if the transparency is applied to the F. 5 plate, we find that it comes much closer to the Tritan confusion lines.



WILLMER W.11

In terms of colour difference the total colour distance from the position of the background to the midpoint for the green square is 33 N.B.S. units, while the distance from the background to the blue square is only 11 N.B.S. units.

The other two Tritan plates used in this research were the Willmer's W.2 and W.11. From the analysis of results obtained on the age population, it was found that W.2 is sensitive to age variations giving an inverted V type of curve for those with clear recognition of the numeral 2, and that after the 6th age group, progressively more and more fail to read this plate, rising to 60 per cent of non-recognition for the last age group.

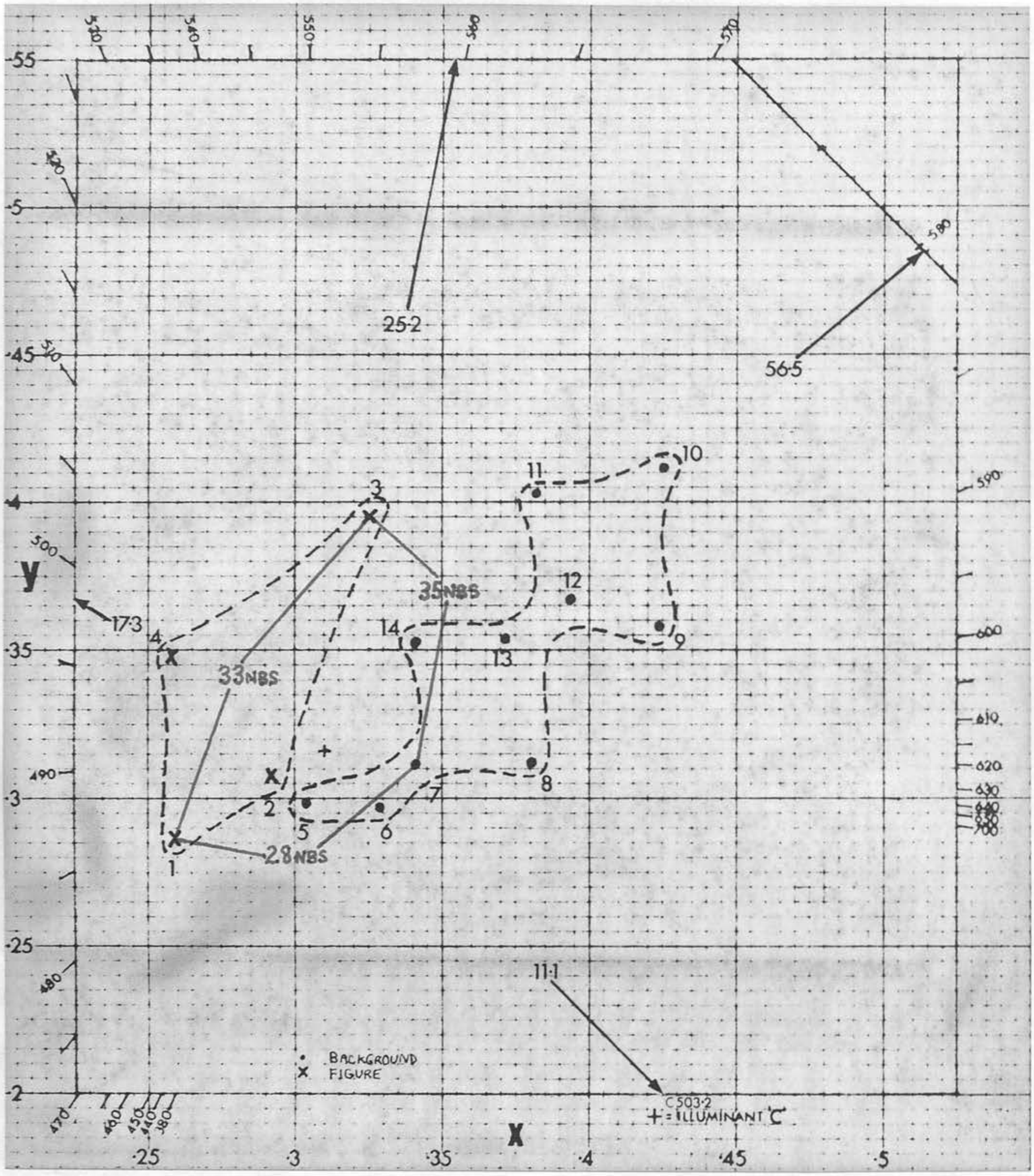
As far as W.11 is concerned, a rather curious distribution was found in the frequency of those would could read this plate at two or three feet (this is the distance at which only Tritan subjects are supposed to read it). Only half of the subjects read it as a clear 11, and fewer subjects (approximately 45 per cent) read it at a distance of five to seven feet. It was concluded from this data that the test plate is not reliable for showing age variations.

The colorimetric and photometric principles involved in good test construction have already been pointed out, and it therefore comes as no surprise that the W.11 plate is an example of how pseudo-isochromatic plates should <u>not</u> be constructed. From the objective measurements made, the picture that emerges is fantastic and one is left wondering exactly what Willmer was trying to achieve. Fourteen different reflectance curves have been recorded for this plate (see the photometric data sheet) four for the figure and ten for the background. No more need be said - a glance at the graph is enough.

The colorimetric data sheet shows that the colours employed in the

GRID No.	Placing	λD (mμ)	pe	CO-ORDINATES x	y	Y%	Average Y%
Figure 1	2:4	484		·2593	·2857	28·46	
2	5:8	'C'		·2921	·3077	43·34	29·60
3	2:3	559·5	25·2	·3251	·3953	30·05	
4	7:4	498	17·3	·2580	·3484	16·58	
Background							
5	6:1	'C'		·3040	·2980	31·72	
6	3:10	c 503·2	11·1	·3288	·2973	26·72	
7	3:5	'C'		·3409	·3109	46·11	
8	8:10	630		·3797	·3117	21·72	
9	1:2	591·0		·4235	·3578	21·84	33·06
10	5:1	580·3	56·5	·4253	·4107	32·84	
11	5:5	575·2		·3822	·4030	46·85	
12	6:6	584·8		·3935	·3672	36·85	
13	1:6	584·8		·3719	·3539	36·07	
14	1:3	575·2		·3414	·3529	30·00	

Table No. — COLORIMETRIC DATA FOR Willmer W. 11.



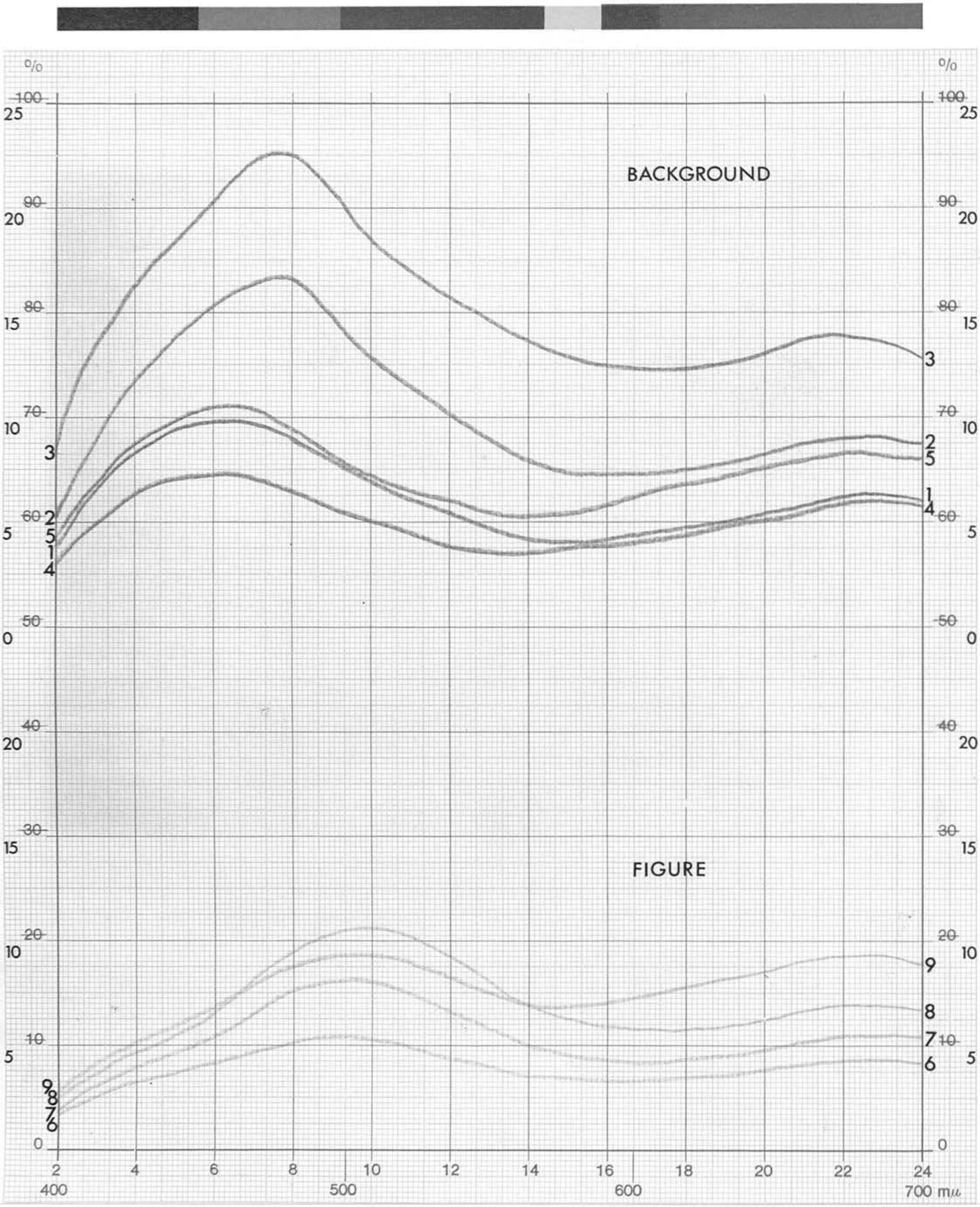
y
x
·55
·5
·45
·4
·35
·3
·25
·2
·25
·3
·35
·4
·45
·5
25·2
56·5
17·3
11·1
33NBS
35NBS
28NBS
C503·2
• BACKGROUND
x FIGURE
+ = ILLUMINANT 'C'

WILLMER W.11

figure are composed of colours whose dominant wavelengths range from 559.5 to 484 mu, while for the background this ranges from c 503 to 530 mu at the green part of the spectrum. On the chromaticity diagram there are two clusters of colours, in each case spreading over a very large area of the colour space. Perhaps the intention was to construct the test so that the colours of the figure and those of the background would fall upon the Tritan confusion lines. However, the individual colours of each cluster (for figure and background) are far too widespread covering between three to five separate Tritan confusion areas, thus frustrating the initial objective.

But what about people with a normal colour vision ? For these subjects, the total colour difference between the peripheral dots of the figure is 33 N. B. S. units, while the difference between the blue of the figure and the purple area of the background is about 28 N. B. S., and about 35 N. B. S. units from this purple to the green colours of the figure. These are rather long distances for 'normal' subjects to confuse. In the Dvorine the most sensitive test plate was shown to be one where the distance was of the order of 10 N. B. S. units. In another respect, that is in terms of hue or wavelength discrimination, figure and background are separated by a minimum distance of 20 mu, and a maximum of 100 mu. But this is speculation, since it is difficult to find what really happens in this test. Perhaps the photographic analysis will provide another more meaningful or more acceptable answer.

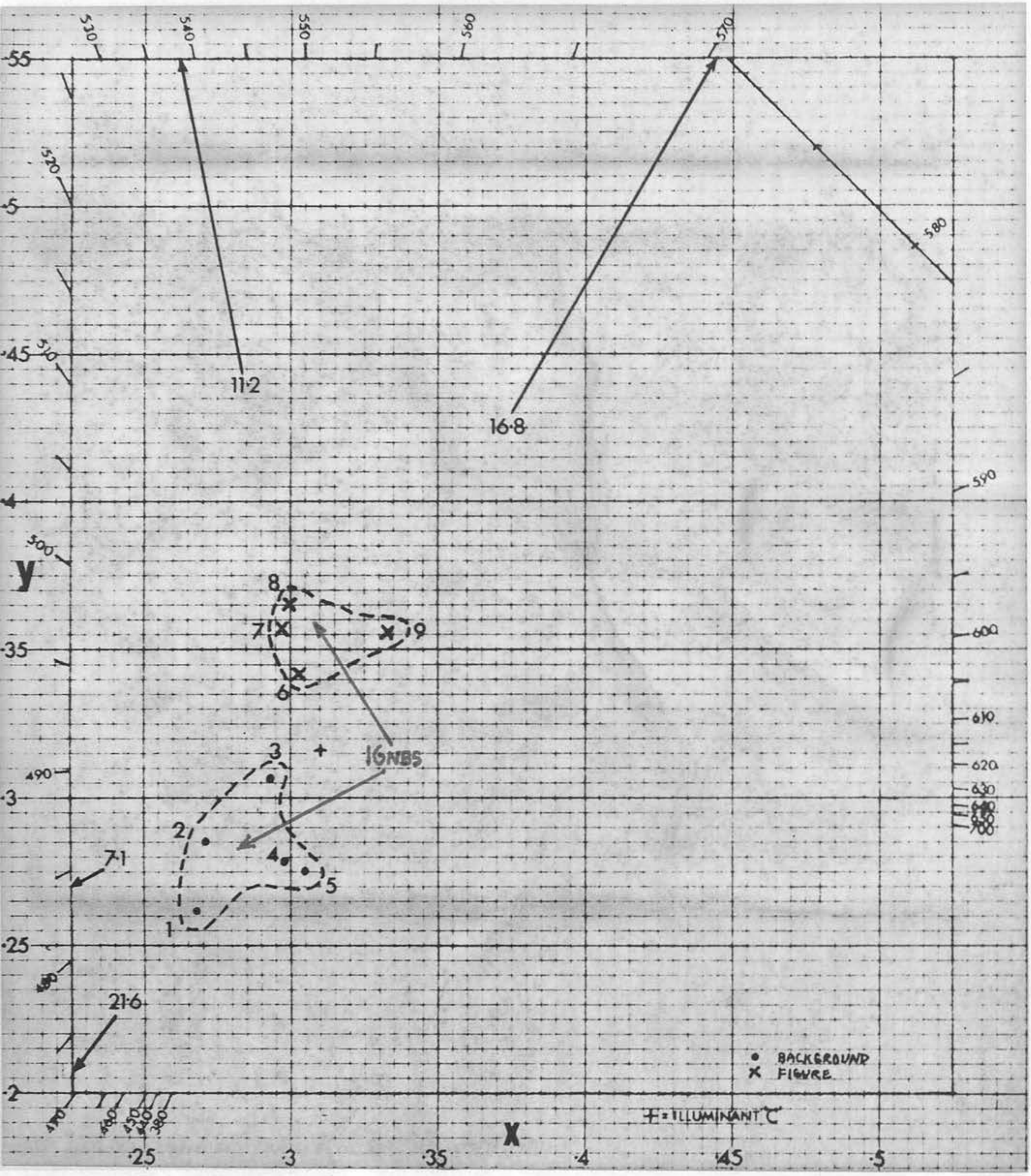
Though the W. 11 is very confusing, the same cannot be said of the other Willmer plate used, namely plate W. 2. The design is basically that of the



WILLMER W.2

GRID No.	GRID Placing	λD (mμ)	pe	CO-ORDINATES x	CO-ORDINATES y	Y%	Average Y%
Figure							
6	13:21	538.5		·3033	·3417		
7	14:23	538·5		·2970	·3569		
8	14:21	538·5	11·2	·2999	·3653		
9	18:18	570·5	16·8	·3328	·3558		
Background							
1	17:20	472	21·6				
2	17:19	480	10·1	·2705	·2850		
3	14:24	C	7·0	·2928	·3059		
4	11:20	c 564·5		·2980	·2780		
5	11:29	c 558		·3049	·2754		

Table No. __ COLORIMETRIC DATA FOR Wilkmor W-2



WILLMER W.2

'vanishing' type (the tritanopes are not supposed to be able to read the number 2) and fewer basic colours have been used. From photometric study four distinct reflectance curves for the figure and five for the background have been identified. The coloured dots 6, 7 and 8 of the figure indicate the use of one ink, but variations in weighting, and their curves, indicate only a small percentage reflectance (in the region of 5 to 10 per cent). The peak for these three colours is at 520 mu. The curve for the fourth colour (No. 9) shows an increase in the red end of the spectrum. The back ground consisting of curves for dots 4, 1 and 5, shows similar characteristics and overall reflectance value again of the order of 5 to 10 per cent. The curves for colour dots No. 2 and 3 show the highest reflectance, up to 15 and 20 per cent at the peak of this curve, that is at 480 mu.

So far, the colorimetric data sheet confirms the inference made. The figure is made of three coloured dots with a dominant wavelength of 538.5 mu and one colour dot at 570 mu, while the background is composed of colour dots in the vicinity of illuminant 'C' some lying in the purple region with complementary wavelength at c 558 to 564 mu, and two dots with their dominant wavelength of 470 and 480 mu respectively.

The chromaticity diagram shows the position of the individual dots in the colour space. Figure and background are placed on either side of the locus for illuminant 'C'. The colour dots are spaced fairly closely around the Tritan confusion lines, and the spread of the colours for the figure and for the background is not more than 8 mu. From the point of view of the 'colour confusion' theory this plate should be a successful and sensitive test for detecting Tritan defects.

Kalmus (1956) confirms this inference. But there is still the question of how sensitive this plate is to minor variations in colour discrimination at the blue end of the spectrum.

Results from the age population study show that this plate is quite sensitive to the age variable. How is this possible, and in what way is it different in this respect from the W. 11 plate ?

By reference to the chromaticity diagram, it is seen that the total colour differences between the mid-point of the figure, and the mid-point of the background is only 16 N.B.S. units, and it must be still less from the nearest colour dot of the figure to the nearest colour dot of the background. Again, if we may draw a parallel, it was found from the analysis of previous tests, that the plates showing the greatest sensitivity to the age variable are found where colour differences are in the region of 10 to 15 N.B.S. units. Plate W. 2 is no exception to this.

IV. Photographic evidence - The colorimetric data sheets for the plates analysed here also gives the luminosity values of the individual colour dots making up either the figure or background. This was deliberately not stressed since it is fully investigated in this section dealing with photograph evidence.

Average luminosity differences between figure and ground are given first for the Ishihara. Only three series will be discussed. For the first series, the average luminosity difference is 7.9%, for the second it is 8.8% and the third 8.3%.

The Spectromat results show the following luminosity values for plates 2 to 5.

Figure		Background	
Orange	Pink	Green	Blue/Green
38.3	-	24.7	29.5
49.1	43.7	42.4	-
56.6	56.8	48.8	59.5

Average luminosity - 48.9%   Average luminosity - 41.0%

Brightness differences are less noticeable on the Dvorine than in the Ishihara. Nonetheless they do exist, but often exist only between two of the four colours used in each plate, the other two being of equal brightness. The probable result will be that a coherent 'readable' pattern cannot be formed as frequently as on the Ishihara.

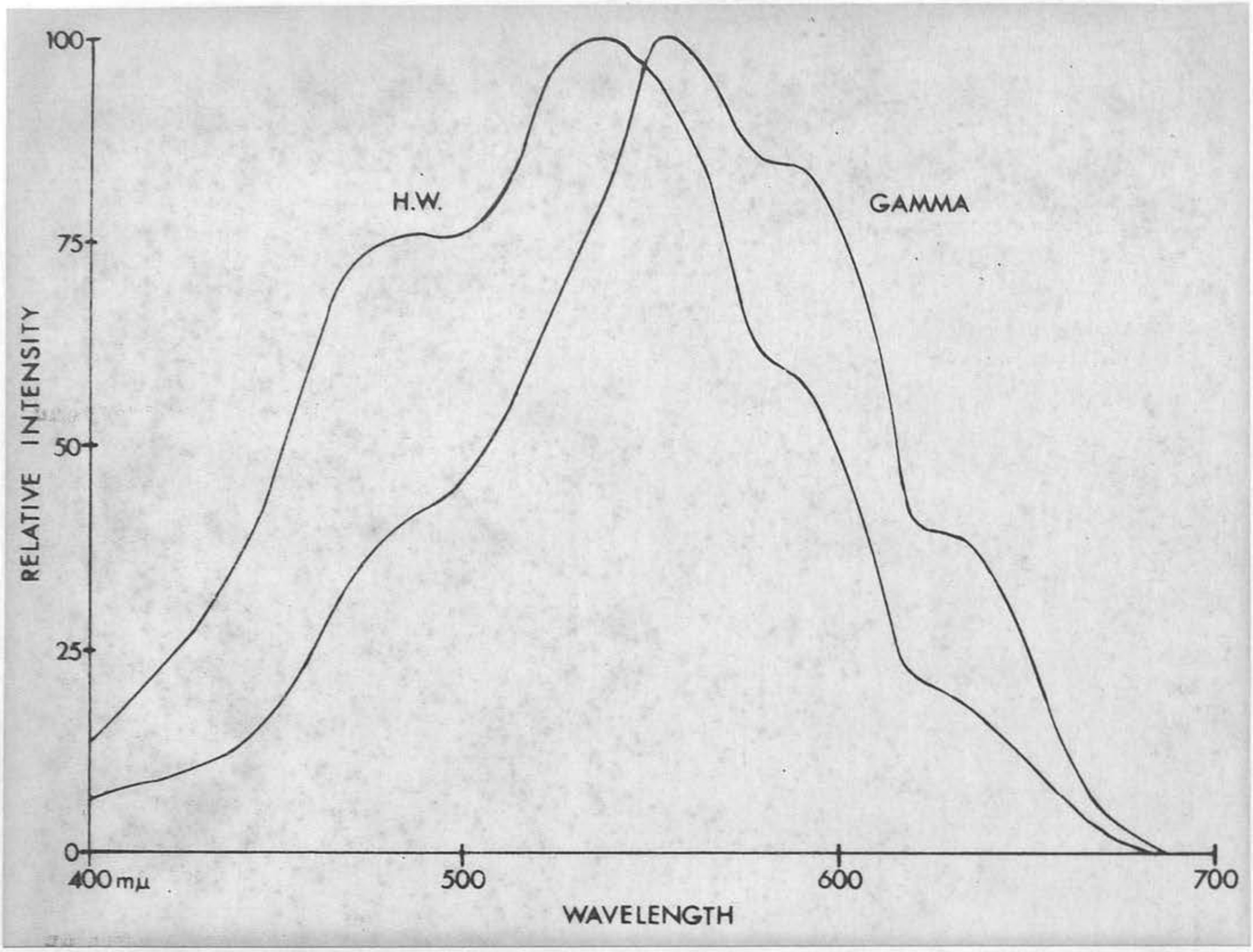
The most serious differences are found in the four plates of series 1 and 11.

	Figure	Background
Plates 2 and 3	Green	Brown
	41.7	33.7
	62.3	65.2
Plates 4 and 5	Figure	Background
	Brown	Red
	35.2	27.5
	35.2	50.2

Those differences of 8 to 10 per cent between the average luminosity values of figure and background could be of great consequence where subjects find difficulty in reading the plates. It might be that the brightness element makes it easier to identify the figure if it helps it to stand out more clearly against the background, so that discrimination of the figures becomes largely dependant, not on perception of hue differences, but rather on brightness differences.

From the evidence so far gathered, it is safe to assume that the age effect leads to a desaturation of spectral colours, and to a neutralizing of the colours near the centre of the chromaticity diagram, thus causing a relatively larger area to be now referred to as 'white'. If this is the case, then luminosity differences in the plates will become relatively more and more important as the process of 'neutralization' progresses with age, and as people become older their discrimination will depend more on brightness differences, than on hue differences. If there is a change in the relative position of the luminosity function for such subjects the percentage luminosity they see should also differ from the percentage luminosity as measured for the Standard Observer, and this would either help or hinder them in reading the pseudo-isochromatic plates. In short, it provides an additional clue.

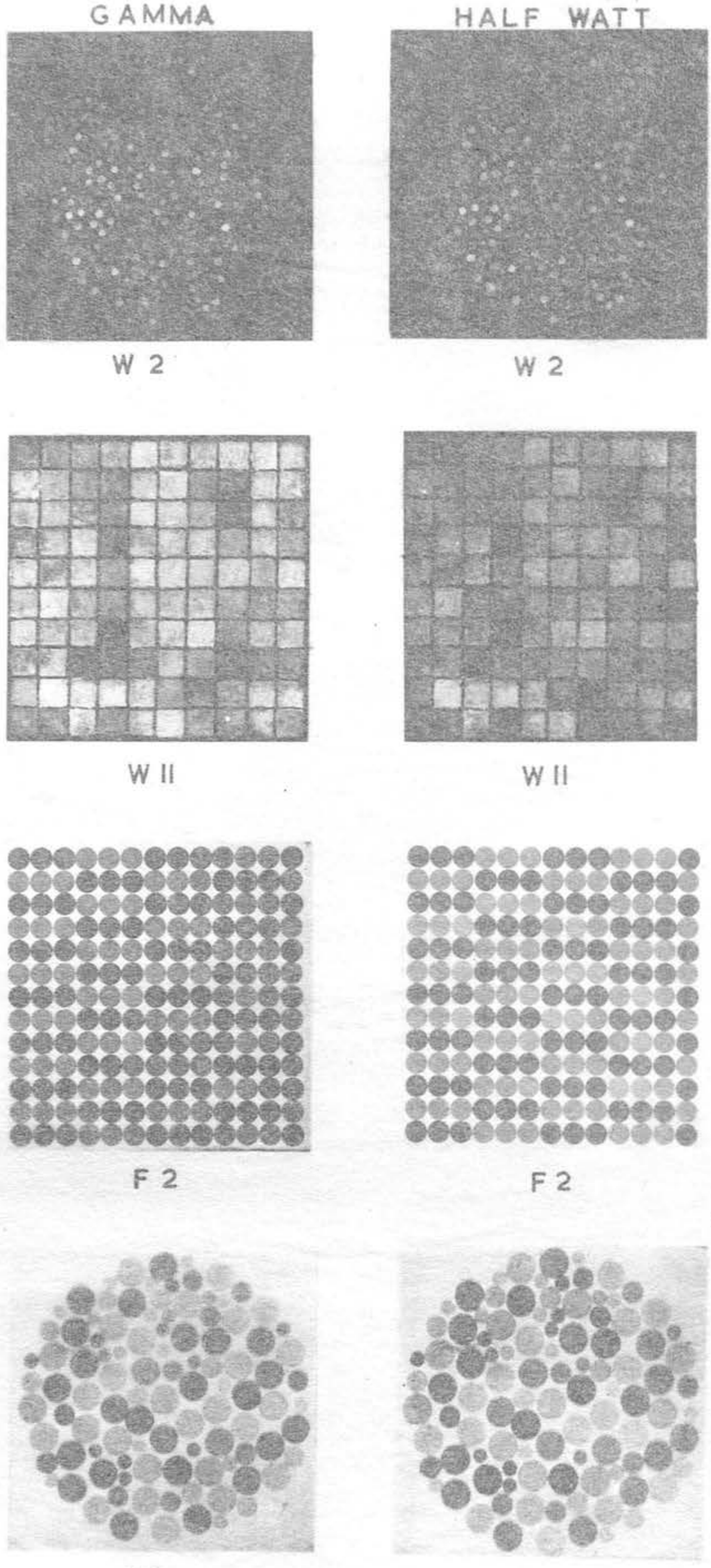
Assuming this to be true, then it would be interesting to see if the pseudo-isochromatic plates possess this additional 'brightness' clue, and to what extent they do so. It is not necessary to find the whole numeral standing out against the background for it to provide this clue. It might only be necessary to see a vague outline for this to afford an additional clue in identifying a particular test plate. As early as 1958, to test this hypothesis, a photographic analysis of the plates was undertaken by the writer but this is the first time that the results have been publicised.



LUMINOSITY CURVES

The plates used in this research were photographed under two conditions, one resembling the luminosity function with a peak at about 555 mm, and the other with a shift to 535 mm. To achieve such luminosity functions for a photographic situation, a laborious analysis had to be undertaken to find the combination of film, filter, developer, and camera set up, which would correspond most closely to the luminosity curves for the photopic function. By using two Ilford filters - Gamma 402 and H. W. 403 - plus Ilford film FP3, and tungsten illuminant A - luminosity curves were obtained for a standard observer, and a typical protan subject. The attempt to find a third curve similar to the deutan's luminosity function had to be abandoned, because of the unsuitability of the colour filters. A graph is presented giving outlines of these two curves for the Gamma 402, and the H. W. 403 combinations. As developing and printing can either enhance or diminish the contrast effect in the final print, it was decided to develop to an arbitrary contrast level of gamma .6. Unfortunately, the initial contrast present in the negatives was distorted during the printing of the plates shown in this thesis. First a master copy had to be made and, though it already shows some distortion, this represents what is on the negatives more faithfully than the second print. The copies finally shown here are reproductions of the master copies and thus, though greater contrast has been obtained for some of the plates, the fine contrast of others has been lost and with it the shadowing figures seen on the master copies.

This method of analysis has pitfalls, but provided that this is remembered, it should be very useful :



TRITAN PLATES

(1) It should now be possible to assess how good the luminosity control of a given test is, and so relative comparisons can be made between tests.

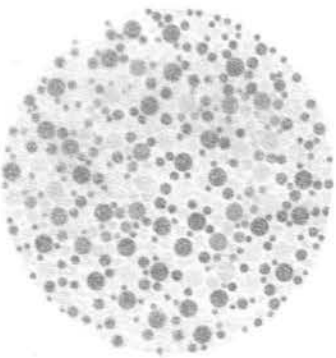
(2) With the two filter analysis method which produces the two luminosity functions under which test has been photographed, it should be possible to find whether changes in an observer's luminosity function can alter the 'luminosity clue' element of the plates.

(i) Tritan Plates - The illustrations of the tritan plates consist of two sets of photographs for each of the four plates. Willmer's plates are marked as W. 2 and W. 11 and Farnsworth's plates as F. 2 and F. 5. At the top of the illustration the name of the relevant filter is found. (Remember that the Gamma filter combination represented luminosity sensitivity with a peak at 555 mu, while the H. W. combination represented one of 535 mu).

This reproduction closely resembles the original plate, and there is neither loss nor increase in the detail presented in these plates. In the W. 2 plates there is some hint of a numeral 2 with both filters and in the W. 11 plate the number '11' is seen with the Gamma filter but less distinctly with the H. W. filter combination.

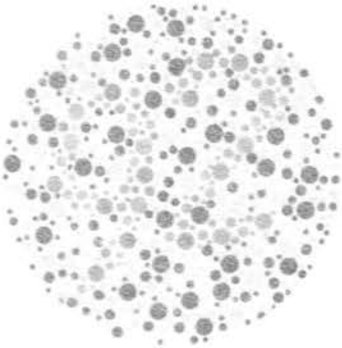
Both Farnsworth plates show neither figure nor pattern and it can be concluded from this that unlike W. 11, control of the luminance factor has been achieved, and the relevant design or number can be distinguished by hue and saturation differences alone.

(ii) Dvorine and Ishihara Plates - The first comparison of the tests using the photographic method is between the prints made with the H. W. filter combination

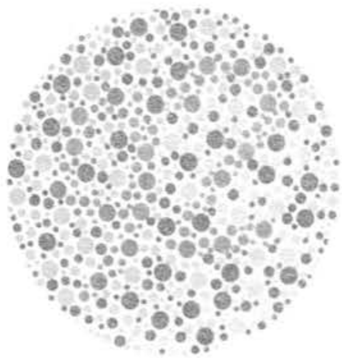


2- 67

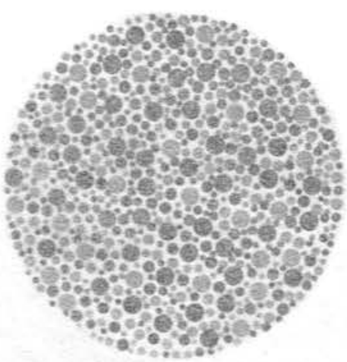
3 - 38



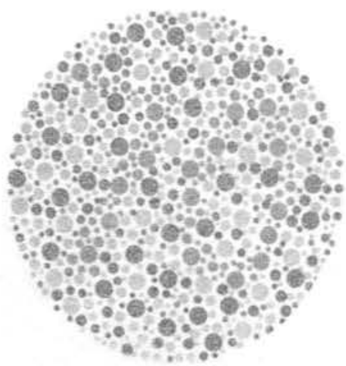
4 - 92



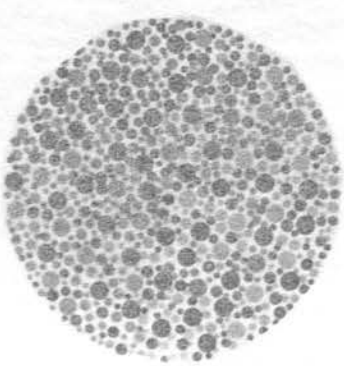
5 - 70



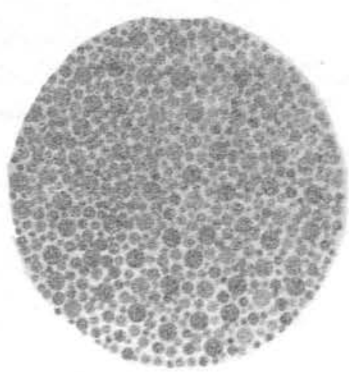
6 - 95



7 - 26

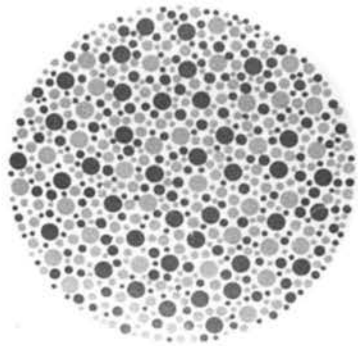


8 - 2

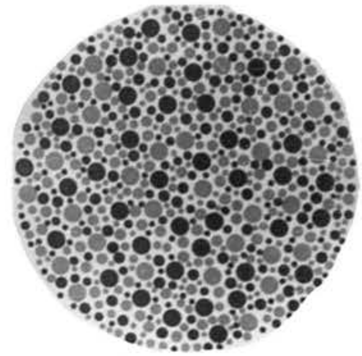


9 - 74

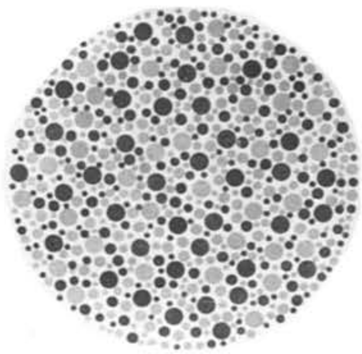
DVORINE H.W.



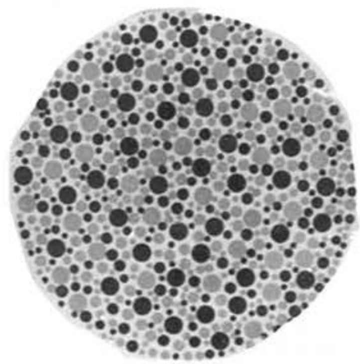
10— 62



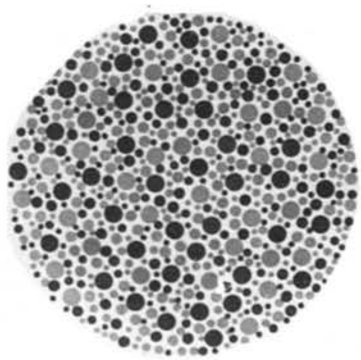
11 — 4



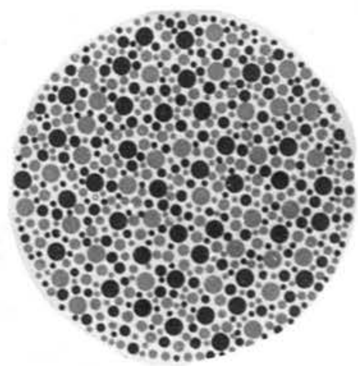
12— 28



13— 46



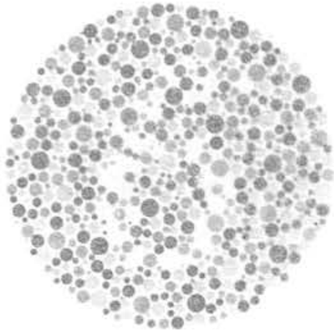
14 — 7



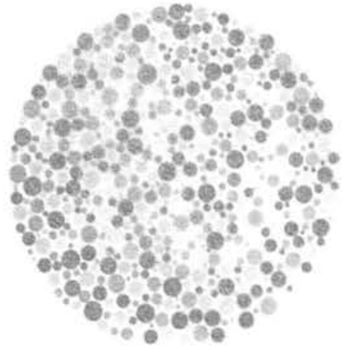
15 — 39

DVORINE H.W.

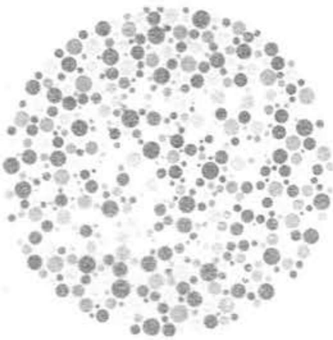
2- 8



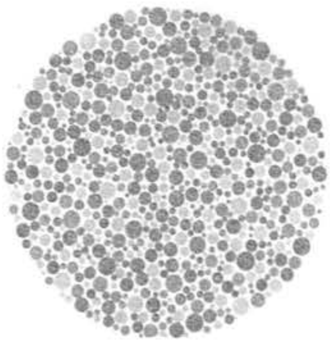
3- 6



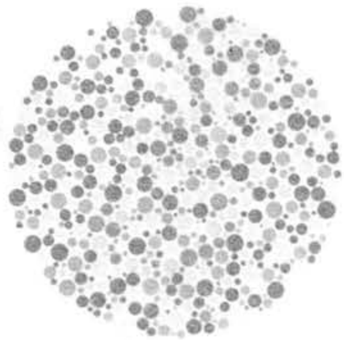
4-29



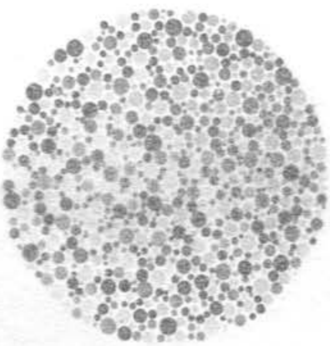
5-57



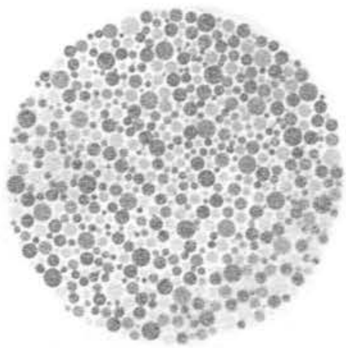
6-5



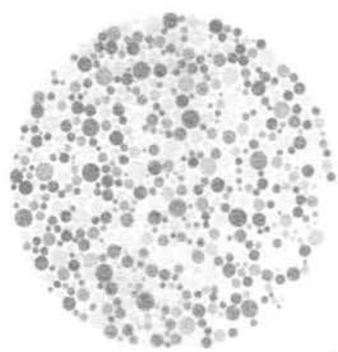
7-3



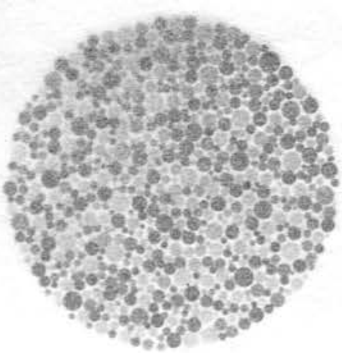
8-15



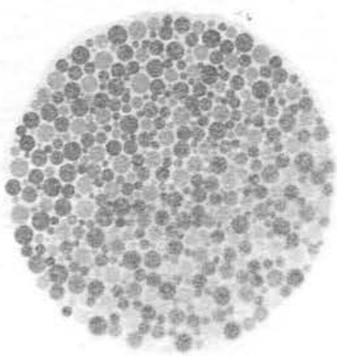
9-74



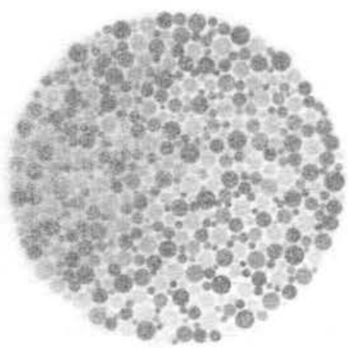
10-2



11-6

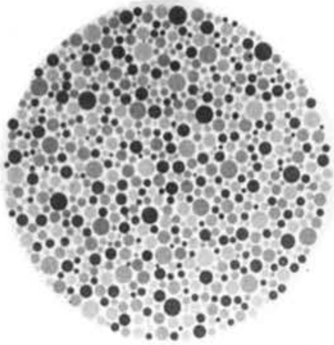


12-97

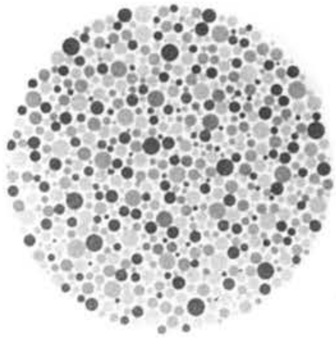


13-45

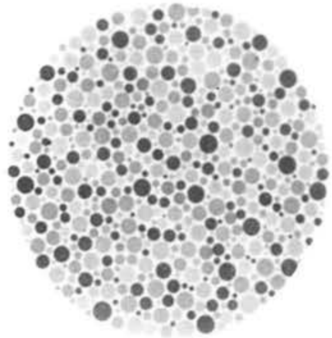
ISHIHARA H.W.



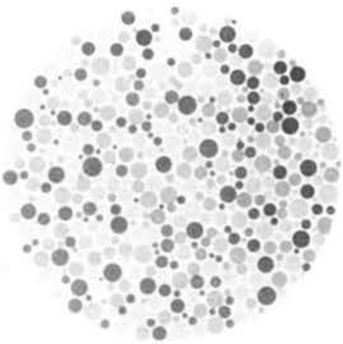
14– 5



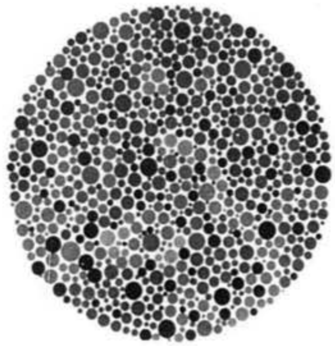
15 – 7



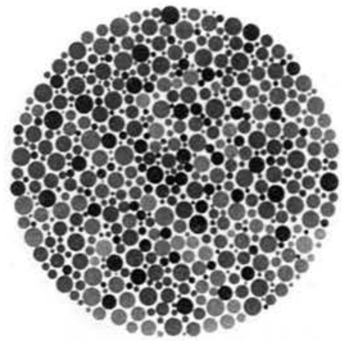
16 – 16



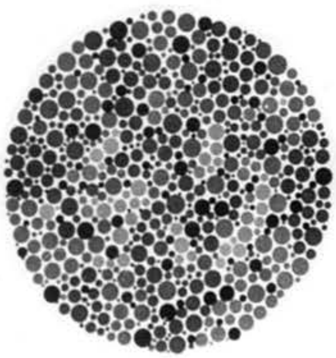
17– 73



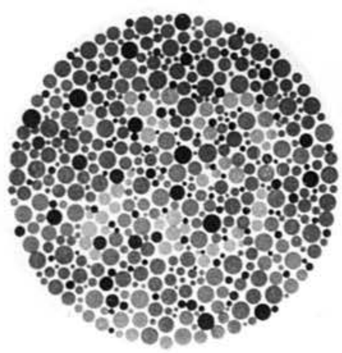
18 –



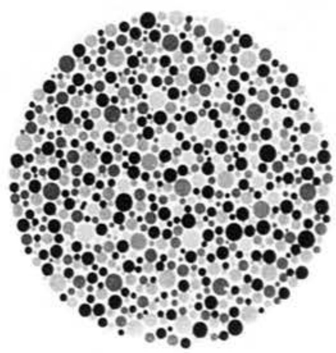
19 –



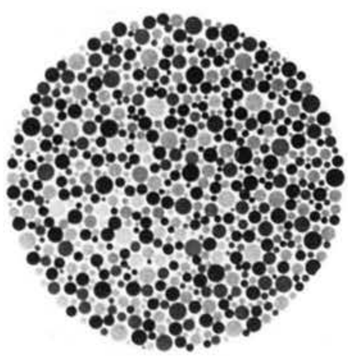
20 –



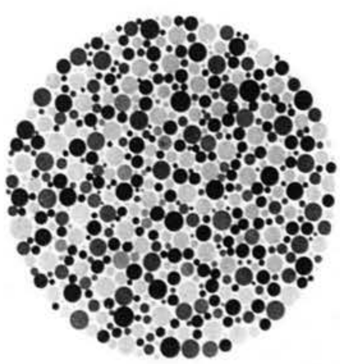
21 –



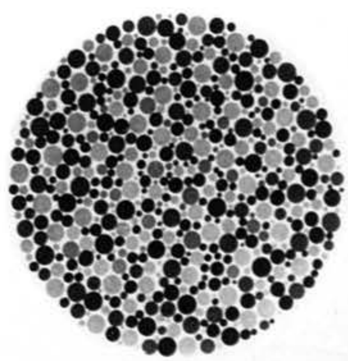
22 – 26



23– 42



24 – 35



25 – 96

ISHIHARA H.W.

(i. e. where the luminosity function has a peak at 535 mu).

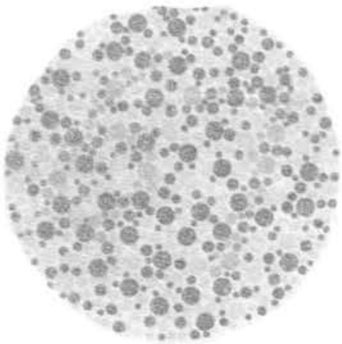
The illustrations here show more contrast than the original prints did, but on the whole, the configurations are the same, when the pseudo-isochromatic plates are photographed in these conditions. Each plate is given its serial number and the numerals a 'normal observer' is supposed to be able to read are quoted.

In the Dvorine, no plate shows figures standing out from the background - though in some there are indications of figures (esp. see plates 2, 4, 5, 6 and 7).

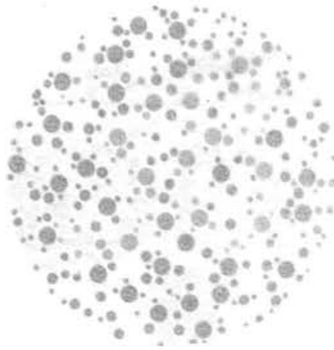
The individual reproductions of the 24 plates of the Ishihara are on two separate prints each with 12 photographs, and again the serial number appears under each along with the numeral that the normal subject would read. There is a slight loss of detail on these reproductions in comparison with the master copy, for example, the numerals in the four plates of the first series are more distinct on the master copy than on the prints appearing here. However, an 8 can be faintly seen, the 6 is clear, there is an outline of the 29, and the 57 is quite distinct. In addition, there is a hint of a numeral 2 in plate 10, and in series 5 plates 20 and 21 show outlines of the 45 and 73.

It is clear that in the series of prints obtained from the H. W. conversion filter six plates in the Ishihara can be read quite distinctly and there are hints of figures in two more, whereas in the Dvorine, no distinct number could be read but a faint outline of the numerals could be seen in five of the plates.

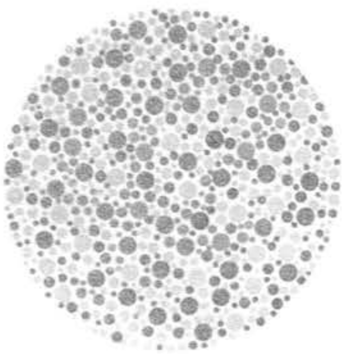
Now let us turn to the photographs obtained using the Gamma filter combination, that is, with a luminosity function showing a peak at 555 mu. (This shows the outline of a luminosity curve for protanopes).



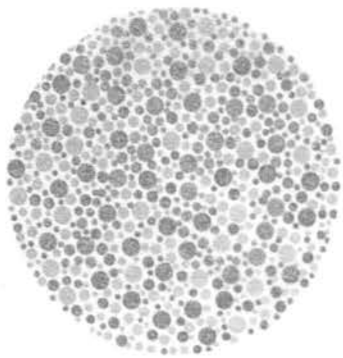
2- 67



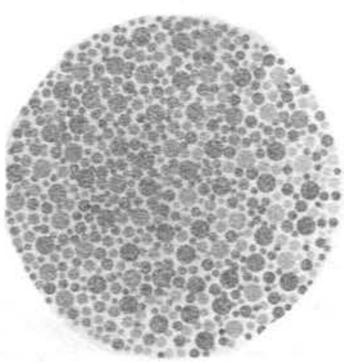
3-38



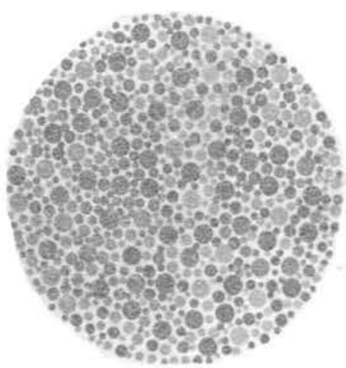
4-92



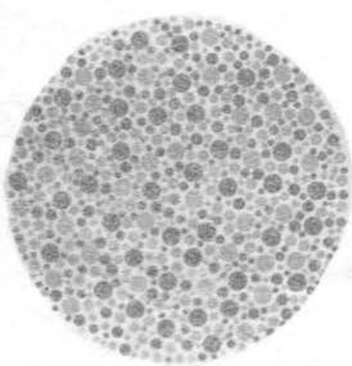
5-70



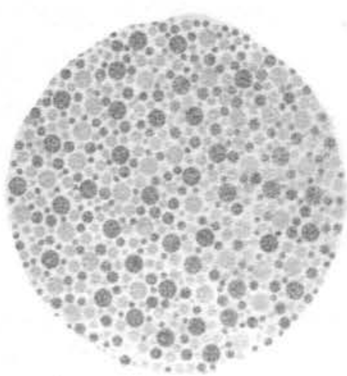
6- 95



7- 26

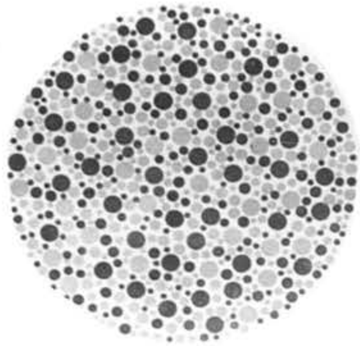


8- 2



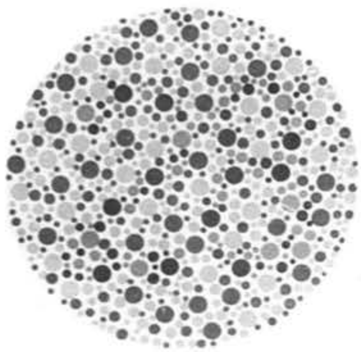
9-74

DVORINE GAMMA

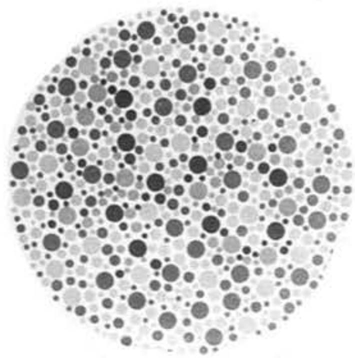


10 – 62

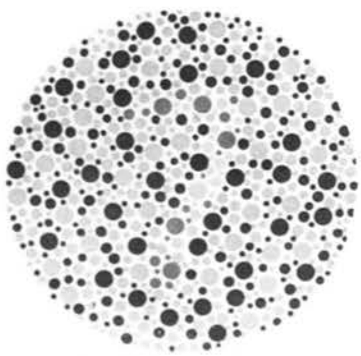
11 – 4



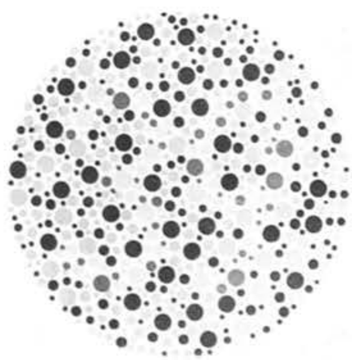
12 – 28



13 – 46

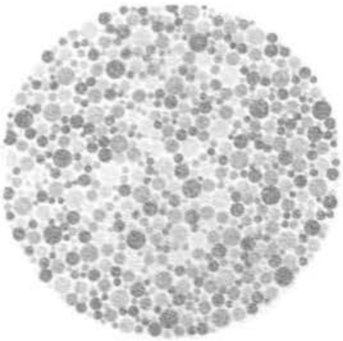


14 – 7

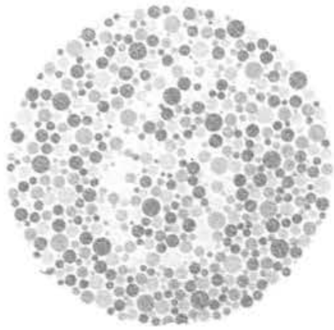


15 – 39

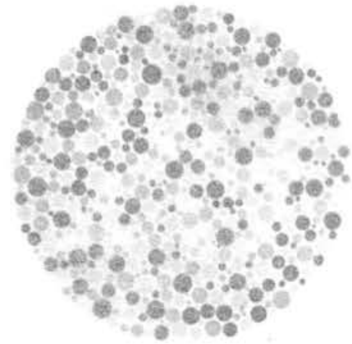
DVORINE GAMMA



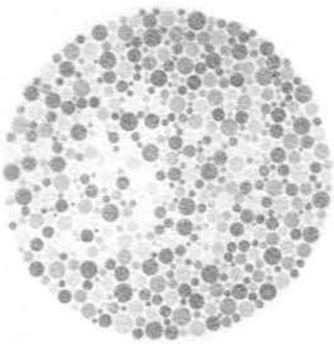
2- 8



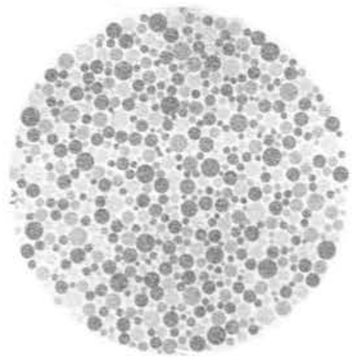
3- 6



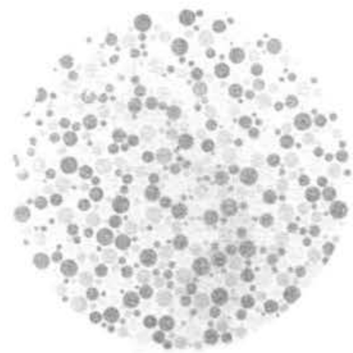
4- 29



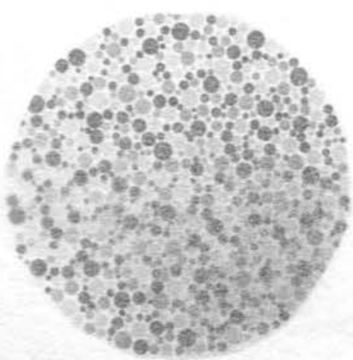
5-57



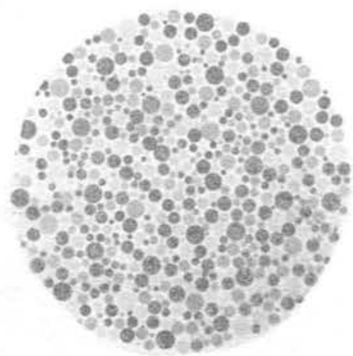
6-5



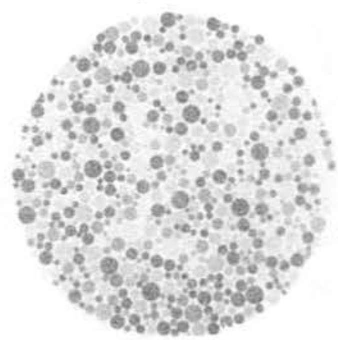
7-3



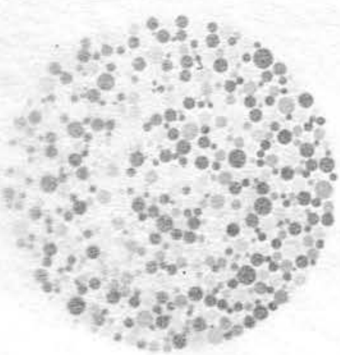
8- 15



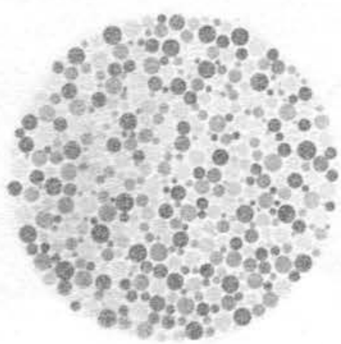
9-74



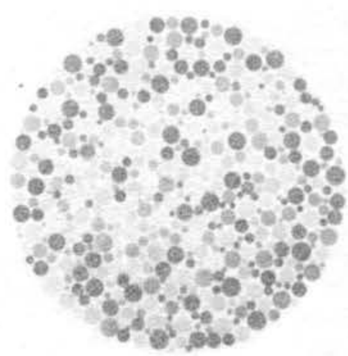
10-2



11- 6

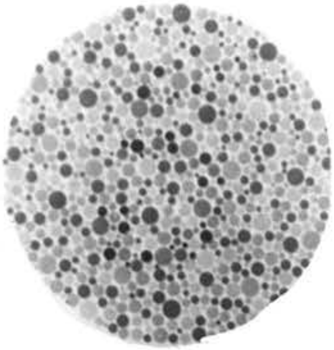


12-97

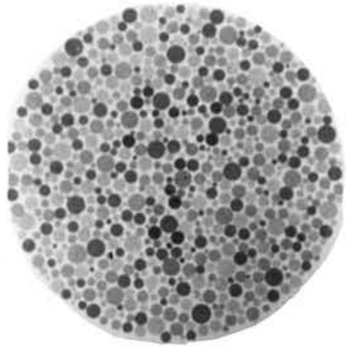


13- 45

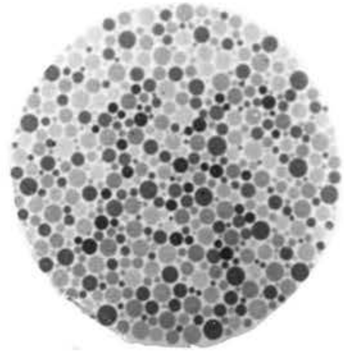
ISHIHARA GAMMA



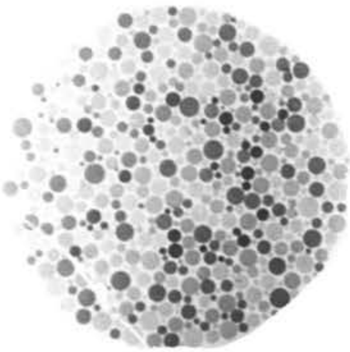
14-5



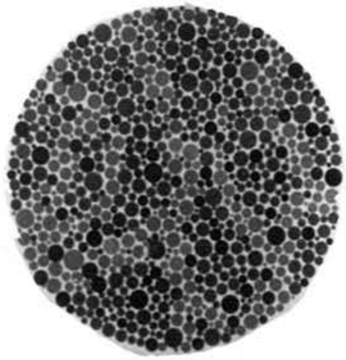
15-7



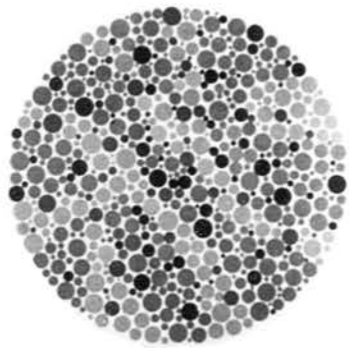
16-16



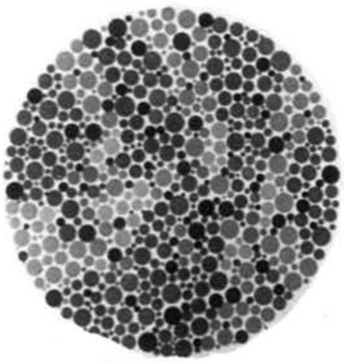
17-73



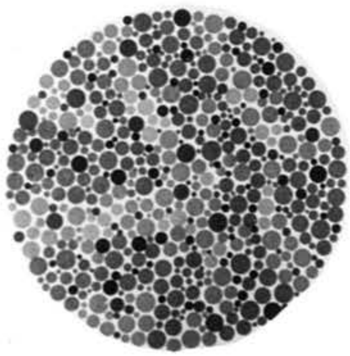
18-



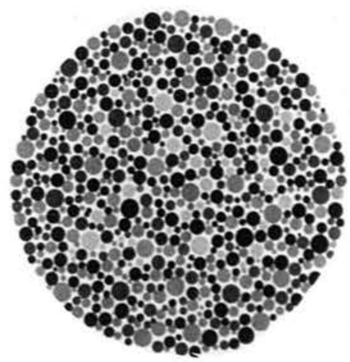
19-



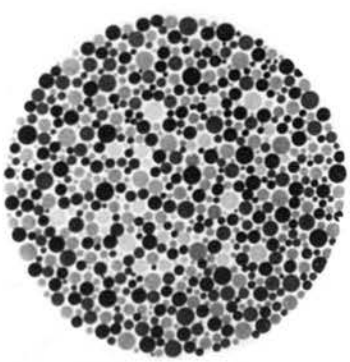
20-



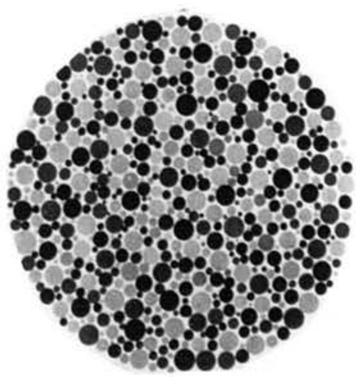
21-



22-26



23- 42



24-35

25-96

ISHIHARA GAMMA

For the Dvorine plates, there seems to be no appreciable difference when using the Gamma filter. Again there is a hint of numerals, but perhaps they are a little more distinct in plate No. 2 reading 67, in plates 4 and 5 and in the two qualitative plates of this test.

In the Ishihara there are interesting changes. Firstly, in the four plates of the series I, plates 2 and 3 are seen more distinctly than before, plate 4 is hardly recognisable, but there is an indication of the 57 on plage 5. In series III plates 10, 11 and 12 show faint numerals, which were not seen with the H. W. filter and in series IV, outlines of the numerals are seen in all four plates, (in plate 14 number 5 and plate 15 number 7, in plate 16 number 16). There is a faint figure in plate 13, but it is difficult to make out what it is and in plate 23 an outline of the 42 can just be detected. There is another difference between the conditions produced by the two different filter combinations. With the H. W. filter, some of the figures in the Hidden Digit series can be seen but no figures can be made out with the Gamma filter. Altogether eight discernable numbers in the Ishihara plates can be read with the help of the Gamma filter and another four to five are partially outlined.

In presenting the analysis and examples above, the aim has never been to discredit the tests but rather to show how carefully pseudo-isochromatic tests must be used. They are not simple tests. On the contrary they provideá complex means of testing Colour Vision where unexpected variables can - and sometimes do - play an important role in how individual plates will be read. Perhaps this factor could explain why some plates such as series I in the Ishihara are rarely misread. How the brightness variable changes with age is not known, but when the Ishihara test in particular, is used (apart from the objective results described in the previous section) the general impression is that older people, say over 40, find the test much easier to do. The brightness clue described provides one possible explanation.

V. Conclusions - In the section on measurements, photometric and colorimetric data was given for the tests used in this research and also for any others that were employed in studies on age and colour vision. It was found that pseudo-isochromatic tests are extremely complex in comparison with, say, the CAT or 100-Hue Test in terms of design and construction. Some test very simple colour functions, while others involve a multiplicity of variables, not necessarily connected with colour.

The measure of the total colour difference gives the best indication of the amount of discrimination involved.

We found that the CAT presented colour problems involving small colour differences of between one and three N.B.S. units, and although such a test might be a very sensitive measure of fine colour discrimination, it is not a good measure of changes in colour discrimination that involve losses of more than three N.B.S. units.

In terms of individual cap differences, the 100-Hue Test works within the same limits as the CAT, but because of the different colour task required of the subject who has to build up colour series over quite large distances on the visible spectrum, it is a measure of both minute and gross losses in colour discrimination.

Plates within even one series of a particular pseudo-isochromatic test vary in difficulty and this is true of all such tests used in this research. Differences in the Ishihara are mostly around 25 to 40 N. B. S. units, although in some plates, the distance between particular colour dots placed at strategic points of a given number might be much smaller than 30 N. B. S. units, in fact they may even be as small as 10 N. B. S. units.

The Dvorine was found to contain the greatest number of plates with figure background distance as small as 12 N. B. S. units, from which we may infer that, in terms of colour discrimination, the test is more sensitive and thus will measure losses of colour discrimination, before the Ishihara will. The qualitatively diagnostic plates in both tests were made in the same manner and it was found that both would have been more effective if the colours chosen for the deutan figure had used more of the available purples.

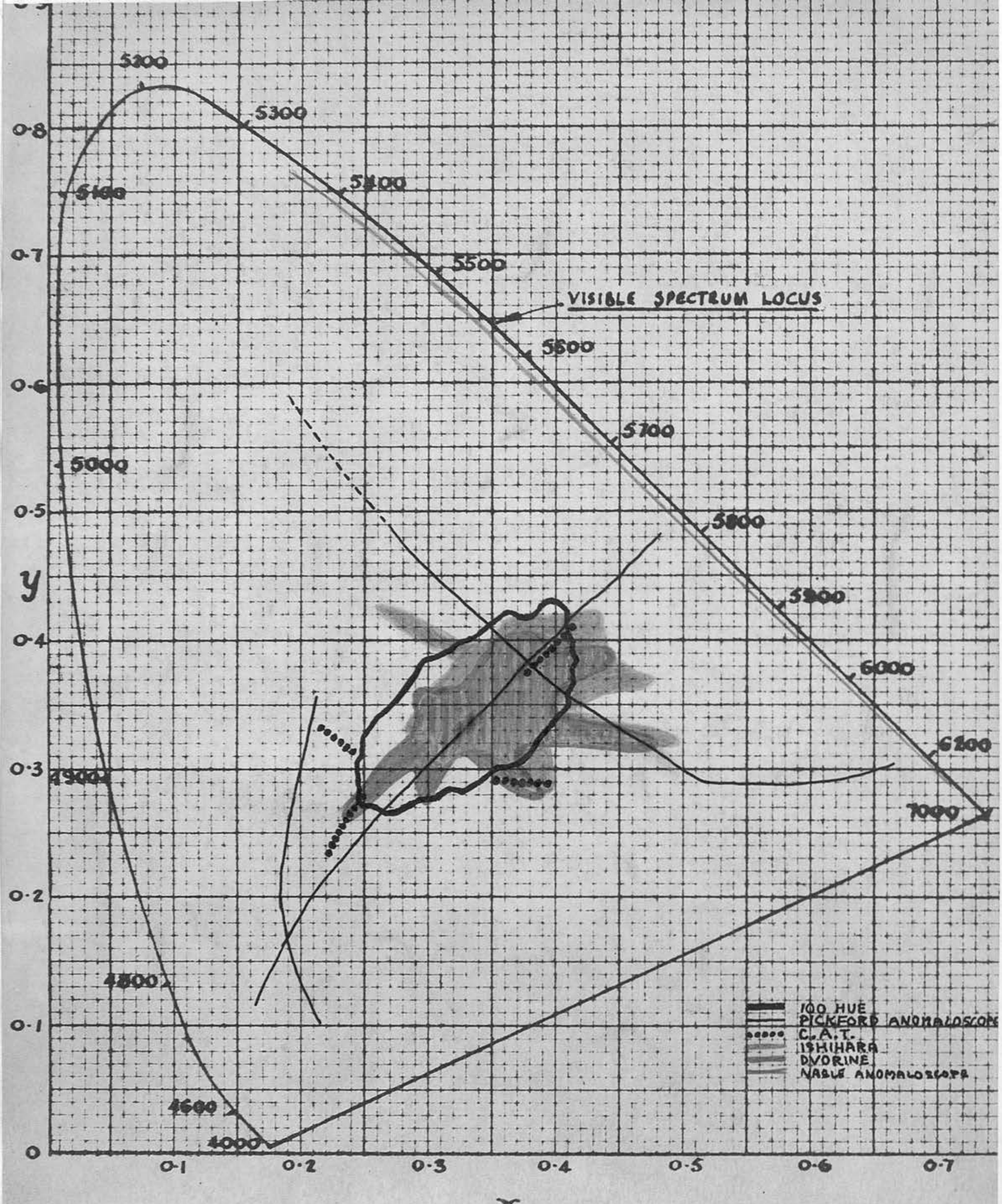
The Tritan plates are of mixed design and some of the less reliable ones showed faulty construction. Again, as in the red-green plates, the total colour differences, exhibited by the Tritan plates, varied from 10 to 40 N. B. S. units.

From the difficulty in perceiving meaningful patterns found in this study of pseudo-isochromatic plates, an interesting hypothesis can be formulated. It may be said that, in terms of colour difference, a rather larger unit than the N. B. S. unit is required if a meaningful colour pattern is to stand out from the colours of the background. It is not enough to be able to perceive the hue of the individual elements in such a field. If the figure is to be clearly recognised as such, the colour differences must be greater than the Just Perceptible Steps that fine discrimination can detect. It is rather like the problem that we come across with the acuity tests. The different elements in a perceptual task can be very small if the task is extremely simple (as, for example, perceiving two separate thin lines as two distinct stimuli) but when the problem becomes more complex (say the type of picture or letter used in common opthalmic acuity charts) in this test situation the absolute differences employed are greater than we could use in a simple acuity test.

In a similar manner, in the simple situation where no meaningful material is used, colour discrimination can be as fine as one Just Perceptible Step (i.e. 1/5th of N.B.S. unit). In plates of the pseudo-isochromatic type the colour differences in absolute terms must be very much greater, say, by at least a factor of 10. To 'perceive' a figure which is made up of discontinuous material (that is, separate dots) the colour difference between the figure and background must be at least 5 N.B.S. units.

In the Dvorine plate No. 10 the total difference is about 12 N.B.S. units. But even with difference of this scale, around 40 per cent in the best age groups cannot 'fuse' the colours into a figure though they were able to name the individual dots correctly according to colour. If the colour distance alters because of increased age and becomes smaller, a point will be reached when it is not large enough for it to be possible to fuse the mass of colour dots into a figure.

Thus whereas in simple hue discrimination tests differences of 6 N.B.S. units can be regarded as rather large, in test situations where other perceptual elements are involved, especially where a meaningful design has to be



5200
5300
5400
5100
5500
VISIBLE SPECTRUM LOCUS
5600
5700
5000
5800
5900
6000
6200
4900
7000
4800
4600
4000
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0
y
0·1
0·2
0·3
0·4
0·5
0·6
0·7
x
100 HUE
PICKFORD ANOMALOSCOPE
C.A.T.
ISHIHARA
DVORINE
NAGLE ANOMALOSCOPE

identified, colour differences of 6 N.B.S. units are extremely small and can only be regarded as large if they approach 30 or 40 N.B.S. units. Plates for dichromats only should be constructed with colour difference of this scale, but plates for detecting minor defects require colour differences of 6 to 15 N.B.S. units.

From what has been said it is evident that indiscriminate comparison of performances on these colour vision tests is futile for they differ in the degree of difficulty that they present to the testee. They might all be able to detect extreme defects, but anything less than extreme defects should be interpreted strictly in relation to what the plates were designed to test. This has to be found by experience for it is questionable whether the makers of these colour vision tests fully understood the tests they designed.

Chromaticity diagram (opposite) outlines the general positions of the tests discussed in this thesis including these of the three equations for the anomaloscope and merits closer study. It is worth noting that the total colour area occupied by the Dvorine seems to be spread more towards the yellow-blue colour axis than the Ishihara. As the effect of ageing is easier to measure on this axis, it is perhaps not surprising that the Dvorine provides us with a more sensitive test of age differences.

3 : 5 THE ANOMALOSCOPE

Marion P. Willis and Dean Farnsworth in their paper on 'Comparative Evaluation of Anomaloscopes', give this bold definition for an Anomaloscope, 'that it is a device for mixing red and green lights in varying proportions to produce a match for yellow light'. Later, however, they add that if such an instrument includes an adjustment for brightness of the yellow field this can then indicate the differences between the two types of dichromats, that is, it differentiates between protanopia and deuteranopia. They go on to say that an anomaloscope could b e designed to mix any two lights to duplicate a third so long as the types of defect could be shown by abnormal colour mixture. This, of course, is what is generally understood by the term anomaloscope.

I prefer to re-define the anomaloscope as an instrument that is essentially a simple colorimeter. This definition fits the Pickford anomaloscope better than Farnsworth's description because in it three sets of mixtures are used, that is, red-green (the classical Rayleigh equation), yellow-blue and violet-blue-green.

In most researches, anomaloscopes are used for a qualitative analysis of dichromats, (that is, to find deuteranopes and protanopes) to detect anomalous trichromats of the deutan and protan type, (i. e. DA and PA) but also to find those with normal discrimination.

However, the anomaloscope is capable of giving more information than this and what is finally obtained will depend, to a large extent, on the way the results are treated. Most of the results from the anomaloscope come in arbitrary units though for some anomaloscopes the relative energy values for the original primary stimuli used are known initially. Even in such cases the resultant mixtures are expressed only as positions on an arbitrary unit scale, usually represented as so many millimetres on a linear scale or as a degree rotation on a circular scale. If a quantitative account is required from the arbitrary units, the usual procedure is to find the most frequent single point accepted by subjects in a random population and to take this as the modal or mid-matching point - provided, of course, that the random population was truly random. In the Rayleigh equation this denotes that so much red energy plus so much green energy add up to a sensation of a standard stimulus which is usually yellow. However, even among normal people, there are great variations in the proportions of the two primary stimuli required to give a sensation equivalent to the standard. These people deviate from the most frequently accepted ratios, and in order to deal with this problem, various mathematical devices are employed to express the measure of their deviations. For example, on the Nagel Anomaloscope, the ratios are expressed as quotients - thus a quotient of 1.5 denotes a degree of deuteranomaly while 0.5 indicates protanomaly, and any intermediate value can be obtained. The more this value deviates from 1 in either direction, the greater is the deviation of a given subject from the normal most frequent match.

This quotient is based on a formula where

$$Q = \frac{N'}{73-N'} \div \frac{N}{73-N}$$

N is equivalent to the position of the left screw for an average normal subject and N' is the position adopted by a deviant subject.

This anomaly quotient is the result of work by Trendelenburg (1929), but unfortunately, this formula gives an account of extreme deviations only. Its weakness lies in the fact that the scale ends used on the Nagel are not necessarily the same from instrument to instrument. To overcome this difficulty, in his anomaloscope studies, Pickford used sigma scores to give a proper emphasis to the mean normal match, where the size of the deviation is expressed in terms of a statistical parameter, namely, the standard deviation. Thus, one can talk about a given match on the anomaloscope falling within 1, 2 or more standard deviations. Here the limits of plus and minus three times the standard deviation, mark the boundaries of the normal distribution, and outside these limits lie the mid-matching points of genetically determined anomalous trichromats.

Whichever way the results are scored, the anomaloscope is an instrument where colour variations can be measured on a continuous scale and, in this sense, it is a far better instrument for detecting colour vision variations than any other tests that have been described so far, but it is a simple colorimeter and has its own limitations. With such instruments as the Helmholtz colour mixing instrument or Wright's colorimeter a more complex mixture can be chosen, though these instruments depend on very complicated optical systems, and, as the viewing is Maxwellian, a high degree of sophistication is required of the subjects in order to overcome the pitfalls of such viewing. In the Pickford anomaloscope the stimuli of the standard and of the matching side appear on a milk glass and the subject does not look into the instrument but at the stimuli side by side.

The first type of information the anomaloscope yields is the mid-matching point. By the position of the subject's mid-matching point in relation to the most frequently chosen mid-matching point we can detect the type of defect that a given subject possesses.

The Mixing of colours is metameric, that is, the subject has to adjust the ratios of the two primaries to achieve a match to the standard, but physically the colours of the match are essentially dissimilar and it is only in terms of sensation that they are alike. Anomaloscopes differ in the degree of metamerism employed - for example, where very narrow spectral transmissions are used, as in the Nagel, the metamerism is very high. However, there are anomaloscopes that use other means such as colour filters with narrow or broad transmissions. It has been said that instruments with narrow spectral transmissions are superior to those with broad band transmission as far as detecting major defectives is concerned. In a major paper by Louise L. Sloan (1949) a comparison was made between the Nagel Anomaloscope and the dichroic filter anomaloscope. Her general conclusions were that the simple filter anomaloscope differentiated the deuteranomalous and the protanomalous type of defect but did not distinguish between defects among dichromats. This, of course, is because no provision was made for adjusting luminosity values in the instrument and this is the essential difference between the Nagel and the dichroic filter instruments. However, in the Pickford Anomaloscope provision is made for adjusting the brightness. In his comparative study of broad band transmission filters of the gelatine type and very narrow transmission filters of the interference type, Pickford found that there was no essential difference in the results obtained by major defectives and no real difference among the normal population. However, there are the results of Willis and Farnsworth on the differences between various instruments they tested, but these are perhaps not so much due to the breadth of transmission or the purity of radiant energy used, but rather to the fact that the various instruments tested in this study employed different subtends at the retina and different luminosity levels. These variables produce more significant differences than are involved in the problem of narrow versus broad band transmission.

The anomaloscope yields other interesting data, more specifically the so-called 'matching range'. This, in a sense, tells us how many different ratios of one energy to another of one given pair of primaries a subject is prepared to accept as similar to the sensation of the standard. Some subjects accept only one ratio and are very sensitive to any deviation from this ratio. In this case the mid-matching point is also their matching range, the only limiting factor in measuring discrimination (because this is what it is) being the limitation imposed by the mechanical and optical construction of the instrument itself. Many subjects accept a number of ratios as equivalent in sensation to the standard (which does not vary atall during the testing time), and such subjects are relatively less sensitive to changes in stimulus value around their mid-matching point - in Pickford's language, they are the 'colour weak'. In this case the anomaloscope is used as an instrument for measuring the

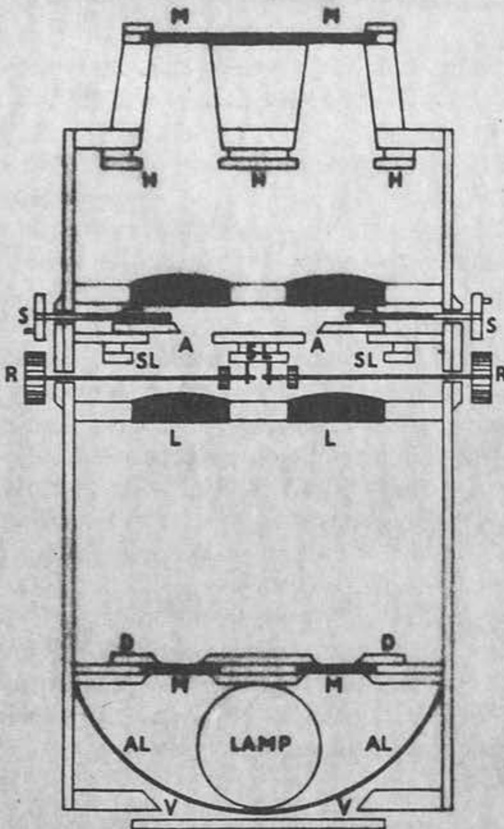


Diagram [redacted]. Plan of the Colorimeter or Anomaloscope

M = Milk Glass Plates
H = Holder for Daylight Filters
S = Shutter Control Screws
R = Knob and Ratchet Slide Controls
SL = Slide Holders
A = Apertures of Lenses
L = Lenses
D = Disk with Apertures for Light Spots
AL = Aluminium Reflector
V = Ventilator

threshold for wavelength change. Many research workers ignore this second type of information as they do not know how to use it - for example where to place the mid-matching point, etc. However, this problem is not peculiar to the anomaloscope, but is a problem in any study of wavelength discrimination. In his book 'Light, Colour and Vision' Le Grand discusses this and how to deal with it when using complicated colorimeters. Partly why this is not dealt with adequately is because there is no real understanding of what this matching range really represents, for after all, there is something very arbitrary in saying that the matching range extends from, say 15 to 35 units, and simply placing the mid-matching value in the centre. This point will be discussed at greater length when measurements are given in terms of colorimetric data for the instrument.

(a) The Pickford Anomaloscope

I. Construction - The anomaloscope used in this study is basically the same instrument as Pickford used for his work on colour vision as described in his book 'Individual Differences in Colour Vision' published in 1949. As the instrument is fully described in the book on pages 1 to 7 and 131 it is only necessary to mention the more technical side of the instrument here, in order to compare it with similar instruments. The energy - that is, luminous flux - for this instrument comes from a 150 watt bulb, which is placed at the back of the instrument and a simple optical system carries this energy across the instrument to the viewing panel. In its path colour filters are placed in a holder that can be moved vertically against the horizontal direction of the light. The problem in constructing anomaloscopes, or any mixing instrument, is to obtain mixtures that give uniform sensation over the entire field of vision. In most instruments that employ Maxwellian viewing this problem is never really solved, as for example, in the Nagel where the primary components of the mixture appear round the fringe of the aperture and where the slightest movement of the head results in a clear perception of them. This becomes quite a nuisance when semi-sophisticated and sophisticated subjects are used, as they always refer to this fact and cannot see the primaries of the mixture as a composite colour or are unable to match the colour with the standard. In more complex instruments, this problem is avoided by the use of an elaborate optical system that is both expensive and difficult to design, build and maintain in working order. In the Pickford anomaloscope this has been solved in a most ingenious way, by using thick glass with a milk surface coating which acts as a scattering agent. Two such surfaces are used, one at the source, the other as a viewing surface. This, of course, solves enormous problems of optical construction, but a heavy penalty has to be paid as these milk glasses are of a density approximating 1 log unit, and as a result of using two such glasses the reduction of luminous intensity is around 2 log units. Thus the loss in brightness is enormous, the resulting intensity being about 1/100th of the original intensity used in this instrument. In this respect the Pickford anomaloscope is a most inefficient instrument. If we now consider that in addition the spectral filters used are also rather dense it will be realised that essentially the test has been reduced to a mesopic level of luminosity, and even verges on the scotopic level. This has certain drawbacks, as subjects being tested have to be dark adapted prior to testing, and this is time-consuming. There is also a doubt about the validity of testing as far as the photopic characteristics of the visual system are concerned. This will be discussed more fully when the photometric calculations are examined.

II. Photometric Measurements - The Schofield Tintometer was used for colorimetric assessment of the position of the various equations in the Maxwellian Colour Triangle. Though it gives photometric data, this could not be utilised as the whole photometric system had to be altered and the transformation tables given for the Tintometer could not be used to convert the arbitrary units to photometric units, so measurements giving luminous intensity of the actual test spot had to be made by means of two photometers - the S.E.I. Exposure photometer and the Schmidt and Haensch Universal photometer. The actual visual measurements were done by a 26 year old observer possessing colour discrimination within $\pm$ Standard Deviation in all three equations. The size of the matching field in the S.E.I. photometer is 2°, but the comparison spot subtends only $\frac{1}{2}$ of a degree, while the size of the aperture on the second photometer (Universal) was adjusted to cover the same area on the retina as was used in the colour vision testing situations.

The results obtained from these measurements represent only tentative values, as accurate measurements were not possible for various reasons, the most important being the effect of simultaneous contrast. This was greatest when measuring the luminosity values of the seven 'primaries' used in the anomaloscope, and was less though still appreciable, for the three mid-matching positions of the equations. The effect of simultaneous contrast upon luminosity judgements is of the order of 0.5 to 1.0 log unit between successive measurements, thus over- and under-estimation can be so great that they might lead to 5 to 10 times larger or smaller units of luminosity being read on the photometers.

Until a flicker photometer is available no more accurate results can be given. Nevertheless the values given now should be a close approximation to the real retinal luminosities.

Thus, colour matching for the modal ratios for the 3 equations are made at the following luminosity levels :-

0.55 cd/m^2 for red-green,

0.30 cd/m^2 for yellow-blue, and

0.20 cd/m^2 for violet-blue/green equations.

The retinal illumination that these luminosity values would represent are approximately :-

- 11.0 trolands for red-green,
- 6.0 trolands for yellow-blue, and
- 2.0 trolands for violet-blue/green equation.

The retinal values will differ from age group to age group, depending on the diameters of the pupils and the extent of pigmentary and other ocular absorptions found in the group.

The results, optimal for a 26 year old observer, indicate that all matches in every equation were made at the mesopic level of retinal illumination.

When comparing the results of this research with those given for the six anomaloscopes studied in Willis and Farnsworth's project, it can be seen that in the anomaloscope used here the effective level of luminosity is comparable only to that employed in the Bausch and Lomb Anomaloscope. All the other anomaloscopes have at least 10 to 50 times more luminosity. However, it is questionable if this comparison is valid as the visual angle subtended for these instruments is between $2\frac{1}{2}$ and 8 degrees while that used on the Pickford for each separate aperture is only 40 minutes. Therefore in this instrument measurements are strictly foveal, whereas the measurements given for the other instruments are essentially extra-foveal or such that they belong to 10 degree photometry.

III. Colorimetric Measurements - Apart from giving ratios or quoting results in terms of sigma deviations from the most frequent score, all results from the anomaloscope can also be expressed in more objective terms provided such measurements are available. To achieve this, it is necessary to measure the various values colorimetrically. This can either be done by purely physical means (i.e. using the photometric curves and then transforming them by using the tristimulus values) or it can be done directly, that is, colorimetrically and this is the method employed here. This method had to be used as no other instrument was available in the department at the time the work was undertaken. It was later found that the amount of energy at the viewing aperture is so small that, in any event, only visual colorimeters are capable of giving a reading and the work therefore had to be done on the Lovibond-Schofield Tintometer where the colour matching was done subjectively by an observer using his eye as a null instrument and where the measuring instrument provided the filters and testing situation (Schofield, R. K. 1939, Tintometer, Salisbury, England.

Directions for the assembly and use of the Lovibond-Schofield Tintometer Type 1a)

This method is reasonably objective. Thus for this anomaloscope, actual colour values at the viewing aperture have been obtained and the arbitrary unit can now be expressed in terms of dominant wavelength change, or in terms of the C.I.E. co-ordinates system by giving the positions in the chromaticity diagram. Lastly, the shifts in matching points or the extent of the matching range can now be expressed either in 'just noticeable steps'[1] as found from Judd's Uniform Chromaticity Diagram or as shifts or ranges in terms of wavelength differences. Such an analysis has great advantages in this study. Firstly, it tells us exactly where and what is being tested and this provided some surprises for us when the results were finally obtained and compared with results arrived at using a priori knowledge of the co-ordinates supplied by the manufacturers, of the Ilford filters.

Secondly, such information allows us to compare the results on this instrument with any other instrument or with the results of other types of colour vision testing material.

Thirdly, the results of these transformations bring out the fact that the anomaloscope scale measurements cannot be regarded as expressing uniform shifts of sensation for each shift in the arbitrary scale unit. However, the uniform shifts hold good for approximately plus or minus twice the sigma shifts around the mid-matching point, but for the end scale values, they cease to be linear. In this thesis results for group performance are given both in terms of arbitrary units and in terms of these new measurements. Thus, wherever diagrams show actual matching ranges of performance in terms of arbitrary units, dominant wavelengths and just noticeable steps are also shown when possible. A complete changeover from the arbitrary unit system to the objective system is desirable but this has not been attempted for two reasons, the first being that although the transformation is as accurate as possible, some doubts about its validity still exist since the amount of work required to do this took approximately four months of constant measurements for the three equations, and though individual checks for each reading have been made, it has not been possible to double check results because of the size of the task.

(i) Modification of the Visual Colorimeter (Tintometer) - Before results for the anomaloscope are listed, the validity of the colorimetric measurements should be considered. The basic problem here is that the luminosity level at the viewing aperture is extremely low. It is within the lower part of the mesopic level and thus any visual measurement becomes exceedingly difficult. Modifications had to be made to the visual colorimeter before they could proceed. The first trial attempts immediately showed that the luminosity level of the visual colorimeter was far too high for measuring the anomaloscope, since by the time the subject was dark-adapted to the spot on the anomaloscope his adaptation was broken by the scattered light from the Tintometer falling upon the viewing aperture of the colorimeter.

The first problem was to reduce the intensity of the measuring instrument yet still be able to measure the colorimetric values of the three equations validly, using the same means. At first an attempt was made to reduce the intensity by building masks in the lower window of the Tintometer (that is, the one where the comparison colour is introduced), so that no stray light from the instrument would fall into the anomaloscope viewing plate. When the first readings were made on the red-green equation interesting results came to light. Contrary to what is expected when very narrow spectral filters are used, the positions for this equation appeared to lie nearer to the centre than to the spectrum loci of the C.I.E. diagram thus showing a highly desaturated equation.

Secondly, when an attempt was made to measure the yellow-blue equation (remember that this equation is much less intense than the red-green) it was found to be impossible because the intensity of the matching arrangement in the Tintometer was too bright. It was no longer a problem of reducing scattered light, but one of reducing the <u>general</u> intensity of the Tintometer sufficiently for it to be comparable with that of the anomaloscope, yet still capable of producing valid measurements.

A first attempt was made using neutral filters. A very dense filter transmitting only 1/36 of the incident light was placed in the Tintometer, this time close to both sources so that both fields were now reduced in intensity by a factor of 36.

However, the suspicion arose that the neutral filter was not so neutral as expected when used at these low intensities. At that time there was no spectrophotometer available to measure the transmission characteristics of this type of filter and therefore an empirical test had to be made. Five plastic samples were first placed in the colorimeter and measured under illuminant 'C' under fully photopic light conditions using a 2° viewing eyepiece. Then the same plastic samples were measured again under illuminant 'C' but this time with the

1/36 neutral filter. The results expressed in the Tintometer 'Red, Yellow and Blue System' and the terms of C.I.E. co-ordinates x, y, z, were computed with the results obtained under the second viewing conditions, and it was found that there were appreciable changes. For 5 plastic samples, pink, biege, green, blue and mauve the average increment in x and y with 1/36 filter were

$$\Delta x = 0.0271 \qquad \Delta y = 0.0227$$

which is approximately 7 times as high as the differences between the individual measurements at photopic level only. The greatest deviation was found for the green and blue plastic samples - which were twice as high as the average value for all the samples.

A new solution had to be sought. This time a wire gauze was fixed at the source of the Tintometer in the pathway of the light, following the practice common in Photometry of reducing luminosity by using copper gauze. The plastic samples were tested again under these conditions. Again the comparison showed that there were important deviations in the values obtained under these luminosity conditions. From these facts, and after an additional study was made of the changes that copper gauze induced on the colour temperature of a source, it was concluded that the effective colour temperature is altered appreciably by copper.

So a new attempt was made. This time steel gauze was used, not bluish steel gauze but one which had been deliberately smoked (i.e. where carbon was deposited on the wire mesh). Yet another attempt was made using special nylon mesh but even this affected the colour temperature. Finally, it was concluded that with the smoked steel wire mesh the results on the plastic samples were comparable to the differences in x, y, z between measurements made just

Red-green equation
1st Attempt
2nd.
3rd
4th
+ = ILLUMINANT 'C'
y
X

within the standard error that was calculated from colour matching at a fully photopic level.

The differences between the measurements for copper and for steel wire mesh for the same 5 plastic samples is as follows :-

Copper gauze	Steel mesh
$\Delta x = 0.0088, \Delta y = 0.0061$	$\Delta x = 0.0025, \Delta y = 0.0026$

The smaller the last digits, the more closely the measurements correspond to each other.

To check this, a separate set of Munsell samples was used, and these were again tested with smoked steel mesh at the photopic and scotopic levels and using a 2^{o} subtend. The results for 6 Munsell samples (chosen at the medium value, and chroma level of 6/6 - 10R, 10GY, 10Y, 10PB, 10B, N6) were found to conform closely to the differences obtained from routine measurements, on the Munsell samples under photopic conditions.

The red-green equation was again measured with a 2^{o} subtend using steel gauze. The second results showed considerable changes, the equation being closer to the spectrum locus and showing a saturation effect of approximately one third on the previous measurement. But when a yellow-blue run through was made it was found that even at this low intensity, light from the Tintometer was still too great as a fair amount of luminous flux was added at the viewing panel of the anomaloscope. So a final change was made by increasing the subtend and then adding masks. A third run on the red-green equation was made. A diagram, on the opposite side, summarises the changes in the position of the loci for the red-green equation during the three trial runs described so far.

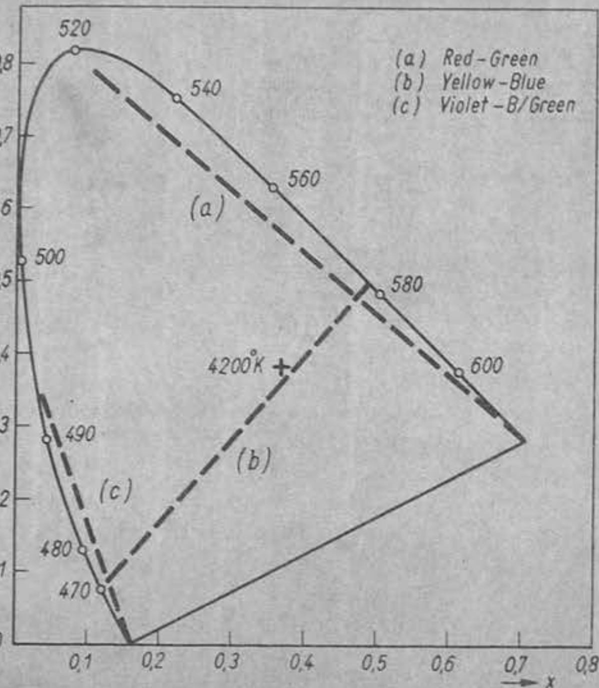


Fig. 1: Position of three colour-matching equations with reference to 4200° K illuminant

Drawn from Ilford specifications

<u>FIND FOLDED C.I.E. AND U.C.S. CHARTS IN</u>

<u>BACK COVER POCKET</u>

(ii) C.I.E. co-ordinates for the three equations - Now routine procedure was adopted using three measurements for each anomaloscope arbitrary unit, each measurement taking approximately 20 mins. to make. The three measurements were averaged and this was taken to be the actual colour value for the arbitrary anomaloscope reading. When the final results were transferred to a C.I.E. chromaticity chart there were only a few irregularities in an otherwise quite smooth graph of the values. These points were measured over again until a correction factor was found. Next, the yellow-blue equation was measured in the same manner, and lastly, the violet-blue-green. The latter caused the greatest difficulty because the level of luminosity was lowest and a complex smoothing had to be introduced. After the actual readings were made the matchings which deviated greatly from the generally accepted curve had to be smoothed out, not empirically on the instrument but by extrapolation on the actual chromaticity diagram.

The results of these measurements have been plotted on two types of diagrams, one on the C.I.E. chromaticity chart, and one on the Uniform Chromaticity Scale devised by Judd. The second chart is useful in measuring the Just Perceptible Difference between points on any of the three equations.

If we take the chromaticity points for the three equations given by Ilford's for the spectrum filters used, and then compare the results obtained from actual measurements, the conclusion that these are two entirely different sets of data is inescapable. The yellow-blue equation is least affected, but the violet-blue-green and the red-green equations show a high degree of desaturation and also shifts in dominant wavelength. The equations are not represented here by straight but by curved lines and the question now is what causes this ? A final answer has not yet been found but there are three possible explanations. Two have been checked and seem to be partly true, but it has not yet been found possible to check the third.

The first problem was to find what caused the desaturation of the primarie used here. It is well known in Photometry that by using milk glass in the viewing field colours can be desaturated, but it is an interesting point that if this type of diffuser is placed between the source and the filter then the filter colour is reduced in luminosity only and not in saturation. This can be demonstrated visually. However, if the milk glass is placed between the filter and the viewer the subjective colour changes and the primary looks very desaturated. This has been tested in two ways - once in the above manner in a visual colorimeter and later an objective test was made on the Spectromat. Results from the instruments illustrate this point (see the following diagrams and two tables of results).

Another explanation for this shift especially in the green and blue-green primaries might be the shifts in the luminosity values at low-level photometry. In his paper on 'Low-Level Luminosity' Weaver (1949) showed extensive changes at this level (i. e. at .1 ft. lamberts or less where the peak of luminance for the first value is now at 545 mu and not 555 mu). This could explain, in part, why these shifts occur.

Then there is still the problem of why the equations are not straight but curved lines. This might be due to an additive failure of extra-foveal vision.

This is not to say that the results are invalid but it should be remembered that the measurement was made using a 3° subtend on the fovea and therefore we could expect some cone-red interaction to take place. The chromatic additivity of tri-variant vision is such that the addition of chromaticity from a mixture of two known stimuli always produces results lying along a straight line. However, when extra-foveal matching is done, the intermediate values between two primaries do not lie in a straight line but form a curve. This has been demonstrated in a recent article on 'Extra-foveal vision in the Additivity concept' by F. J. J. Clark (1962) where his experiments showed conclusively that marked failures of both chromatic and photometric additivity occur for small fields in extra-foveal vision at mesopic levels with an initially dark-adapted retina.

These considerations have a very important bearing on the meaning and relevance of the results in this research. The essential question which has to be considered related to this point. Do the results collected in this thesis represent foveal function or do they present extra-foveal or para-foveal results ? Though the test stimuli in the anomaloscope only subtend 40 mins. for each of the circular apertures, it is not known whether, when looking at the milk glass plate the subject fuses these two, giving him a $1\frac{1}{3}^{\circ}$ subtend, or whether he looks at one first and then compares the other spot (thus introducing a brief time lag). Whichever way it is viewed, the test patch is foveal and does not exceed $1\frac{1}{2}$ degrees. Now in older people, as the losses in luminosity must be taken into account because of the various changes in the ocular media and the retina, we might finally reach a stage when the stimulus level is not sufficient to elicit a response from the foveal region and therefore a para-foveal response alone might result. Unfortunately, it is very difficult to find out whether this happens or not. Only when dealing with older subjects, on a few occasions a subject would find that if he looked straight at the two spots the hue was completely lost, yet remarked that if he looked sideways the colour could be seen. In these cases subjects were advised to look centrally and ignore the peripheral vision.

(iii) Validity of Mesopic Measurements - Because of low-level photometry and the very small angular subtend used, it is important to establish whether the results obtained on the anomaloscope are valid or not. The essential question is whether vision was still tri-variant in the conditions under which testing was done. Le Grand (1957) claims that colour vision remains tri-variant at all photopic levels of luminosities, and, provided that the eye is dark-adapted and vision is strictly foveal, then the standard C.I.E. luminosity curve is accepted as sufficient. What then happens at the mesopic level ? As early as 1891 Abney and Festing, and later Dow in 1906, said that as long as retinal illumination is 0.2 equivalent lux (i.e. about 2 trolands) it can be accepted as tri-variant. Others such as Ives (1912) said that 1 to 10 trolands is adequate illumination and later Rosenburg (1928) and Johnson (1937) said that a retinal illumination of a few trolands is sufficient to elicit a retinal response when the field is 2^{o} in diameter. A larger field of three degrees or more, when viewed laterally, favours a premature appearance of the Purkinje effect. For researches in the field of physiological optics the International Commission on Optics in 1950 recommended the following luminance values and field sizes, for mesopic vision; 0.05 candelas per square metre and 20 degrees viewing. The retinal illumination in the Pickford anomaloscope is in the order of 2 to 11 trolands, and thus, the first objection can be disregarded.

The next question is whether colorimetry is possible at low luminance levels. It seems that because of the independence of the luminous and chromatic phenomena, the latter remains unaltered at least to a first approximation at low luminance levels. Quite early Helmholtz (1896) found that lights remained complementary even when the Purkinje effect completely altered the luminance ratio. Wright (1944) has shown that when the luminance level is divided by 10, colour matches are not altered although the luminous unit of the red primary is reduced to half in comparison with the green. It is therefore possible to assume that the chromaticity co-ordinates of all lights are unvarying and that only the luminous efficiency curve is altered, resulting in the alteration of the luminous units because of the distribution co-efficients. But is it possible to retain a photopic chromaticity diagram and assume a scotopic luminous efficiency function ? Since experiments show that the photopic diagram remains valid to a close approximation, it must be concluded that colour vision at low luminance levels retains the photopic luminosity efficiency but that a supplementary luminance with scotopic luminosity efficiency, (luminance which progressively drowns all colour vision), is superimposed upon it. From the mathematical point of view the tristimulus values of X, Y and Z of the classical C.I.E. system can be retained, but instead of the luminous flux 'F' being identical with Y, (which, with a constant factor can be written $Y = \Sigma V_\lambda E_\lambda \Delta\lambda$), this flux now includes two terms $F = Y + Y'$.

where $Y' = \Sigma V'_\lambda E_\lambda \Delta\lambda$

Y' being the scotopic luminosity factor and 'a' factor which is 0 for photopic vision and increases as the eye adapts itself to darkness. Hence, vision remains essentially tri-variant in each state of adaptation although to present visual magnitudes in all states of adaptation four variables would be needed - two chromatic, and two luminous.

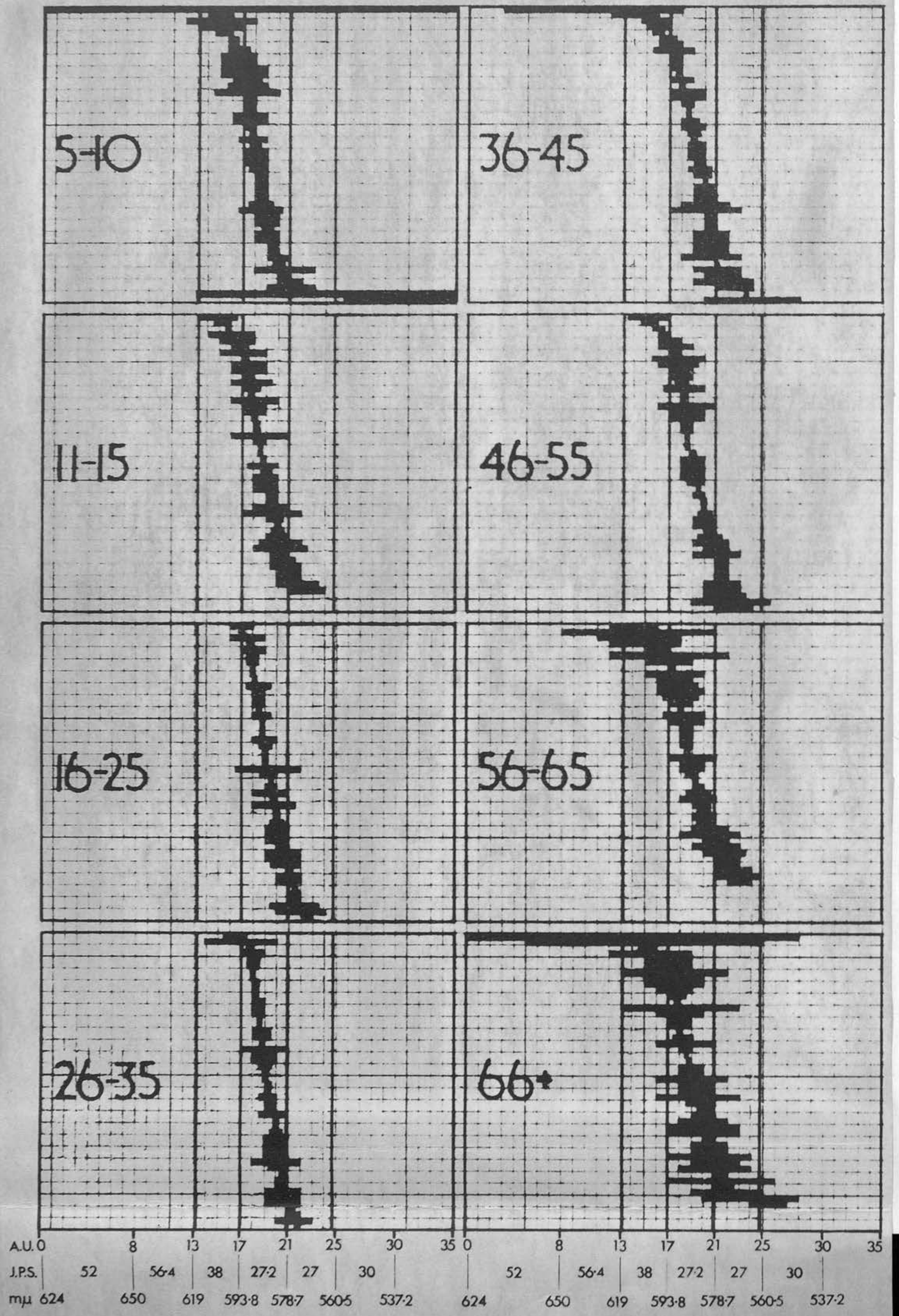
There is also the problem of the variation of saturation with luminance levels.

According to the measurements made by Purdy (1931) for a field of 3°, the maximum saturation for the blue end of the spectrum is attained at a luminance level of about 10 trolands, for the middle spectrum it is about 200 trolands, and for the red end of the spectrum at about 40 trolands.

Perhaps this would explain partly why in the three anomaloscope equations the primaries red, blue and yellow show less of the effects of this desaturation, whereas, the green and green/blue show quite a lot of desaturation.

In addition, Aubert (1865) showed that lowering the saturation not only reduced purity but also produced changes of hue. However, as recently as 1943 Newhall, Nickerson and Judd (with 40 observers) made a study in connection with the Munsell Atlas, where they showed that lines of constant hue do not coincide with straight lines of constant dominant wavelength in the chromaticity diagram but have different curves. These curves vary under different levels of luminance. The specimens with low reflectance factors show a greater curvature than specimens with high reflectance factors.

RED—GREEN


5-10
11-15
16-25
26-35
36-45
46-55
56-65
66+
A.U. 0 8 13 17 21 25 30 35
J.P.S. 52 56·4 38 27·2 27 30
mμ 624 650 619 593·8 578·7 560·5 537·2

(b) Analysis of results obtained on the anomaloscope

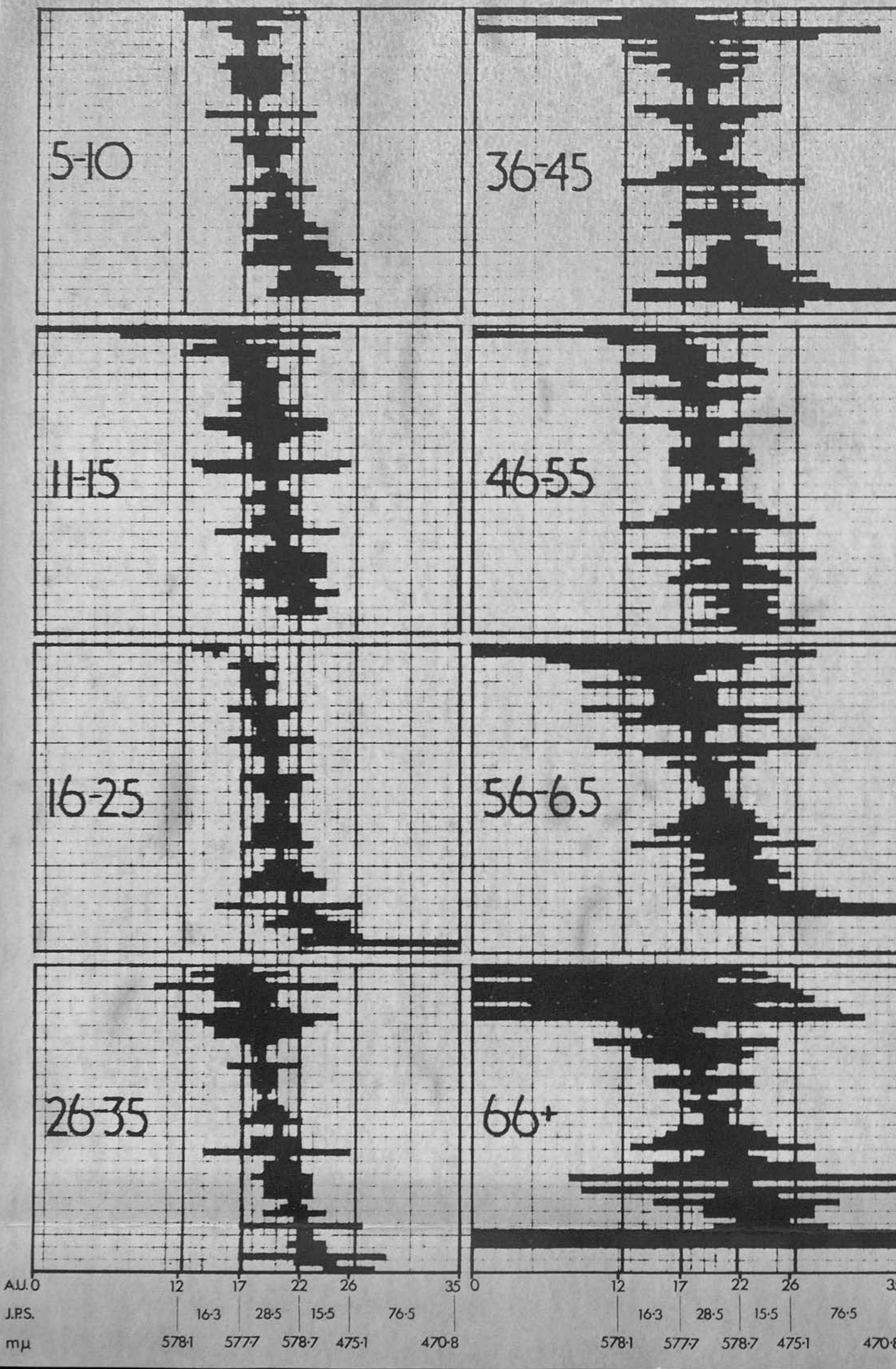
I. Raw Scores - First of all the data is presented in the form of three diagrams showing the raw scores for the red-green, yellow-blue and violet-blue/green equations, and the actual matching ranges for all subjects tested on the anomaloscope for this present analysis. These diagrams are sub-divided into the eight age groups and each result can be read from this diagram in terms of arbitrary units, in terms of Just Perceptible Steps, and in terms of dominant wavelength. To make this more meaningful, population limits were outlined. Thus, the two inner lines in the centre of each diagram represent plus or minus sigma limits, while the two outer lines represent plus or minus three times the standard deviations. Below the diagram, three classes of information are given - firstly, the positions of the arbitrary units on the anomaloscope, then the Just Perceptible Steps values, and lastly, the dominant wavelengths for the various arbitrary loci on the equation.

As previously stated, results were treated in the same manner as Pickford's results, that is, in terms of population parameters but not in terms of anomaloscope quotients. However, when the actual values for the Just Perceptible Steps and wavelength differences are examined, it is immediately obvious that the three scales are hardly correlated at all. If we assume linearity in the arbitrary scale, it is found that this does not hold good for the scaled Just Perceptible Steps within the Uniform Chromaticity Diagram, nor does it hold good for the wavelength increments.

Let us take an example from the red-green equation. Units 17 to 21 on the arbitrary scale represent 27.2 units on the Just Perceptible Steps scale, and 15.1 mu in terms of wavelength difference. In other words, 4 steps on the arbitrary scale represent 27.2 Just Perceptible Steps and 15.1 mu at the most frequent ratios accepted by all populations. Now let us look at the preceding and succeeding 4 arbitrary units (i.e. from say 21 to 25 and then from 17 to 13). It might be expected that because the number of arbitrary units in the second example is the same as the first, the distance on the 'Just Perceptible' scale and in terms of wavelength will be the same. But is this so ? Let us first examine the 4 arbitrary units between 13 and 17. It will be seen from the table that this now represents 38 'Just Perceptible Steps' which is 11 j.p.s more than for the anomaloscope units between 17 to 21, and in terms of wavelength discriminations, this shift represents 25.2 mu which is more than one and a half times the original difference. If we now take the deviation at the green end of the equation (i.e. from 21 to 25) it will be seen that the 'Just Perceptible' differences here cover 27 steps which is the same as the value for 17 to 21, and in terms of wavelength this constitutes 18.2 mu that is, slightly more than the value from 17 to 21. If this example is extended still further, and another 4 steps examined, say from 13 to 8 on the arbitrary scale, it is found that this distance represents 56 j.p.s, while an equal distance on the anomaloscope from 21 to 30 units represents only 30 j.p.s.

We are now presented with the problem of which of the scales to use. If the arbitrary scale unit is used, there is always the impression that this does not represent the real measure of colour discrimination. If the Just Perceptible Steps scale of the Uniform Chromaticity Diagram is chosen, there is the difficulty

YELLOW–BLUE


5-10
11-15
16-25
26-35
36-45
46-55
56-65
66+
A.U. 0 12 17 22 26 35
J.P.S. 16·3 28·5 15·5 76·5
mμ 578·1 577·7 578·7 475·1 470·8
0 12 17 22 26 35
16·3 28·5 15·5 76·5
578·1 577·7 578·7 475·1 470·8

of which should be used. Though it is rather old-fashioned, Judd's system was used because this was available for our direct readings on the Tintometer, and no other was found to be suitable. This system was formulated in 1935 and there are certain objections to its colour space, especially in the violet or purple area of the C.I.E. triangle. The other objection to Judd's formula is that it does not include the luminosity factor in its calculations, but this would not be very important in this research for it is assumed that only hue matches are made on the anomaloscope, as the luminosity differences are adjusted for every arbitrary position on the scale. If the aim were to be objective and physical units only were used, results would then have to be placed on the wavelength scale, but hue discrimination is not even proportionately related to this variable.

It is very tempting to present these results on the Uniform Chromaticity Scale. For the purpose of this thesis, however, this method was rejected, not because there were objections, but rather because of certain reservations. Though in terms of dominant wavelength and colour space, this measurement is a major contribution to our knowledge and understanding of anomaloscopes and of the Pickford anomaloscope in particular, and though infinite care has been taken in gathering this information (it took over 4 months to collect), results were only obtained on the measurements of one observer. They could not be verified because of difficulty in finding a suitable observer to do this rather tedious work. The writer could not confirm the results himself because his colour vision is defective and could not be used for comparison purposes. Another reason for not transforming all the data into the Uniform Chromaticity System was that the anomaloscope works at critically low luminosity levels -

so low that no margin is left for further checks with objective instruments such as the Recording Spectrophotometer.

However, these objections do not detract from the value of this work.

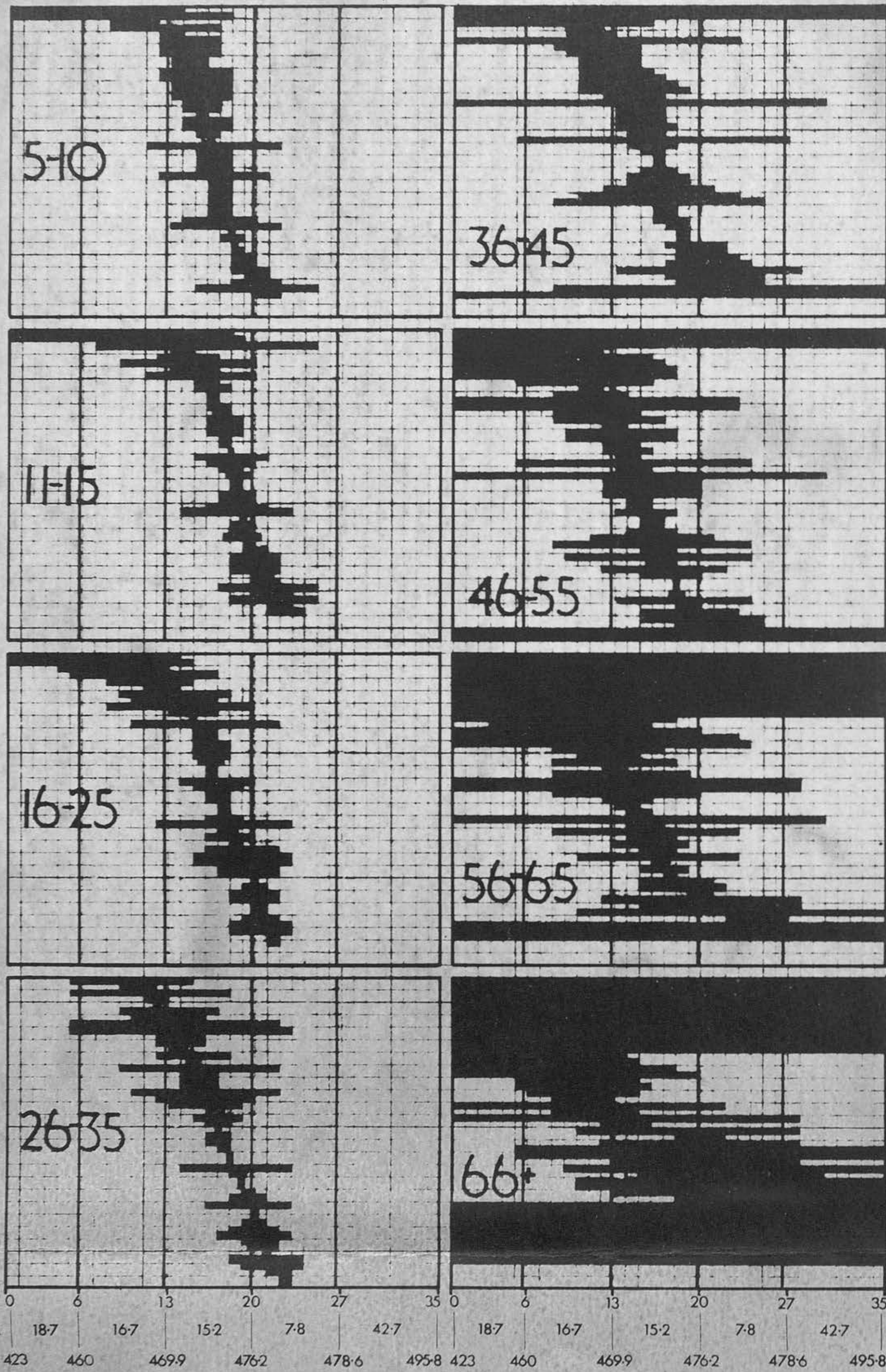
(1) It has produced the very interesting fact that equal steps on the arbitrary unit scale cannot be assumed to represent equal steps in sensation.

(2) As a direct result of this information it is difficult to accept any of Willis' or Farnsworth's correction formulas for transforming anomaloscope results for comparison purposes. If the primaries or the position of the primaries of the various anomaloscopes differ, then the individual steps will differ also, although it is true to say that the results for the most frequently chosen ratio (that is, within I. S. D.), form a linear relationship. Outside this region anything is possible, and unless an empirical study of the particular anomaloscope is made, it is futile to theorise and calculate from a priori knowledge how the space will be arranged for a given anomaloscope.

(3) It is now possible to compare the results obtained on the anomaloscope with results of other colour vision tests, by reference to the actual position each test occupies in the colour space.

There is one final qualification. It is very difficult to determine what one Just Perceptible Step in Judd's Uniform Chromaticity Scale is in relation to the now accepted N. B. S. unit scale. At present, one unit in the new scale is accepted as equivalent to 5 j. p. s., but it is not easy to find whether this applies in Judd's 1935 system. Roughly speaking, if we first take the distance covered by the red-green equation, that is, 270 j. p. s., and then divide this by 5, the

VIOLET—BLUE/GREEN

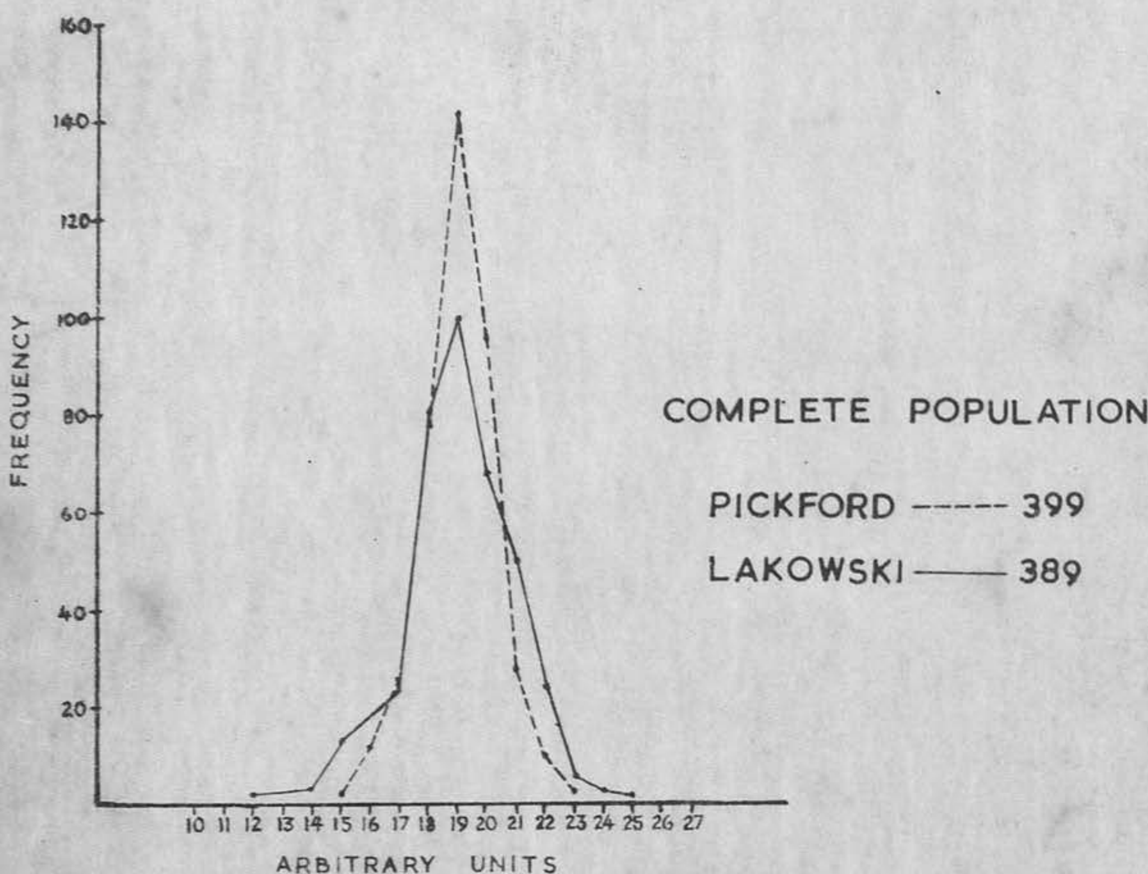


distance for the red-green will then be 54 N. B. S. units. This distance (or colour difference) expressed in N. B. S. units between the extreme points of the red-green equation is quite reasonable, if it is remembered that only hue differences are measured here.

Referring back to the actual charts, it is found from visual inspection that in terms of the distribution of the matches on the three equations, those for the violet-blue/green equation have the greatest spread throughout the equation, and many more people have long matching ranges in all the age groups than were found in either the red-green or yellow-blue equations. It is therefore concluded that in the general population losses are far greater for this equation than they are in the other two. But the differences are rather exaggerated when they are scaled on the arbitrary scale of the anomaloscope. It should be remembered that in terms of j. p. s. steps or even in wavelength distance the violet-blue/green equation covers a shorter distance in the colour space than the red-green equation does. The latter equation covers about 270 j. p. s. and 104 mu, while the former occupies only 101 j. p. s. and some 70 mu in the colour space and so the 35 arbitrary units of the anomaloscope represent a shorter distance than they do in the red-green equation. Let us compare the differences between the limits of plus or minus the standard deviation for both equations. In the red-green equation, the distance is 27 j. p. s. or 15 mu, while in the violet-blue/green equation it is 15. 2 j. p. s. or 6. 3 mu, about a half of the value for red-green. So we see that the visual picture obtained from direct comparison of the two graphs is an exaggerated one, but there are other factors to be considered which point to a difference in

RED-GREEN EQUATION

FREQUENCY DISTRIBUTION OF MID-MATCHING POINTS



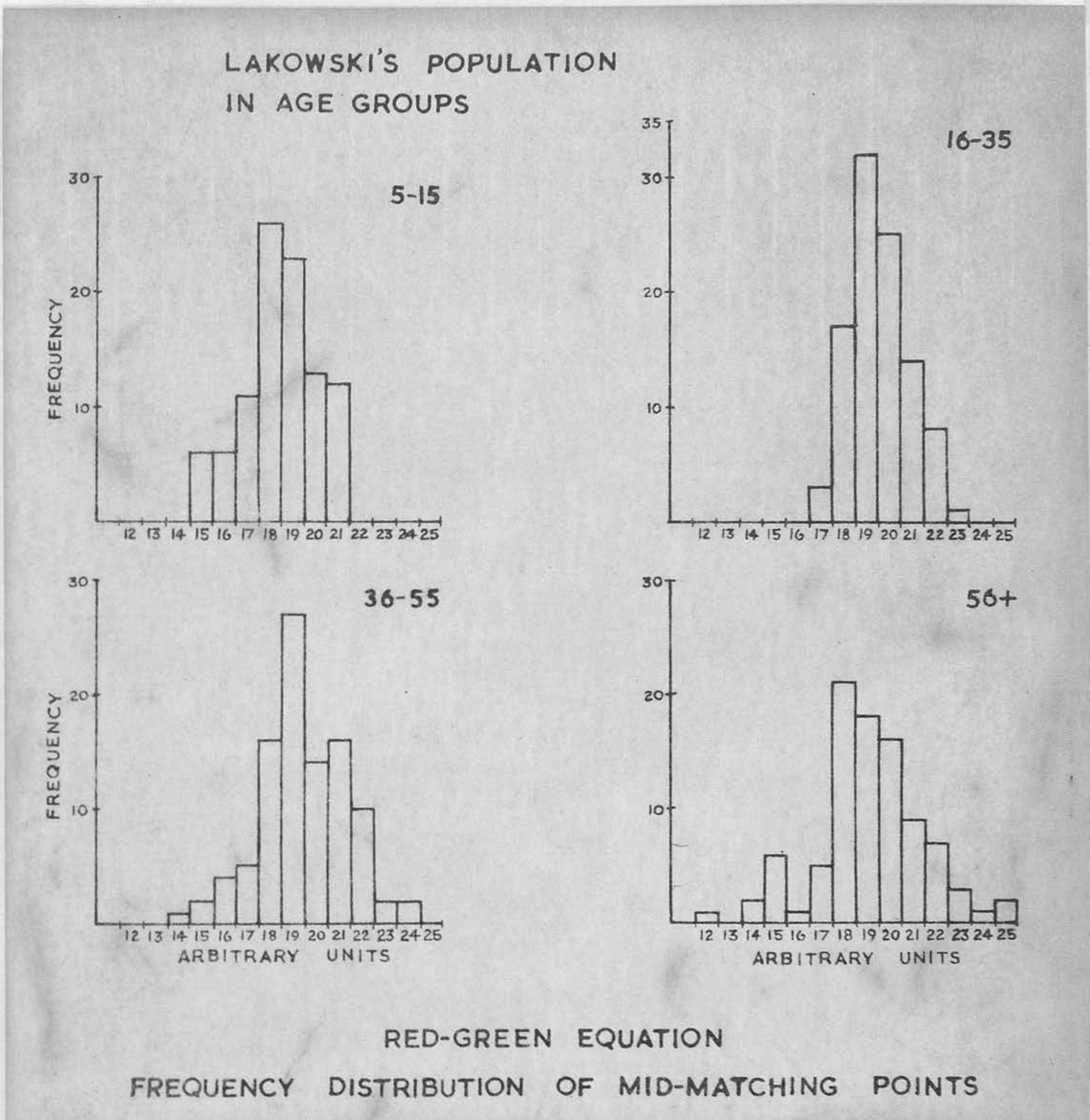
RED-GREEN EQUATION

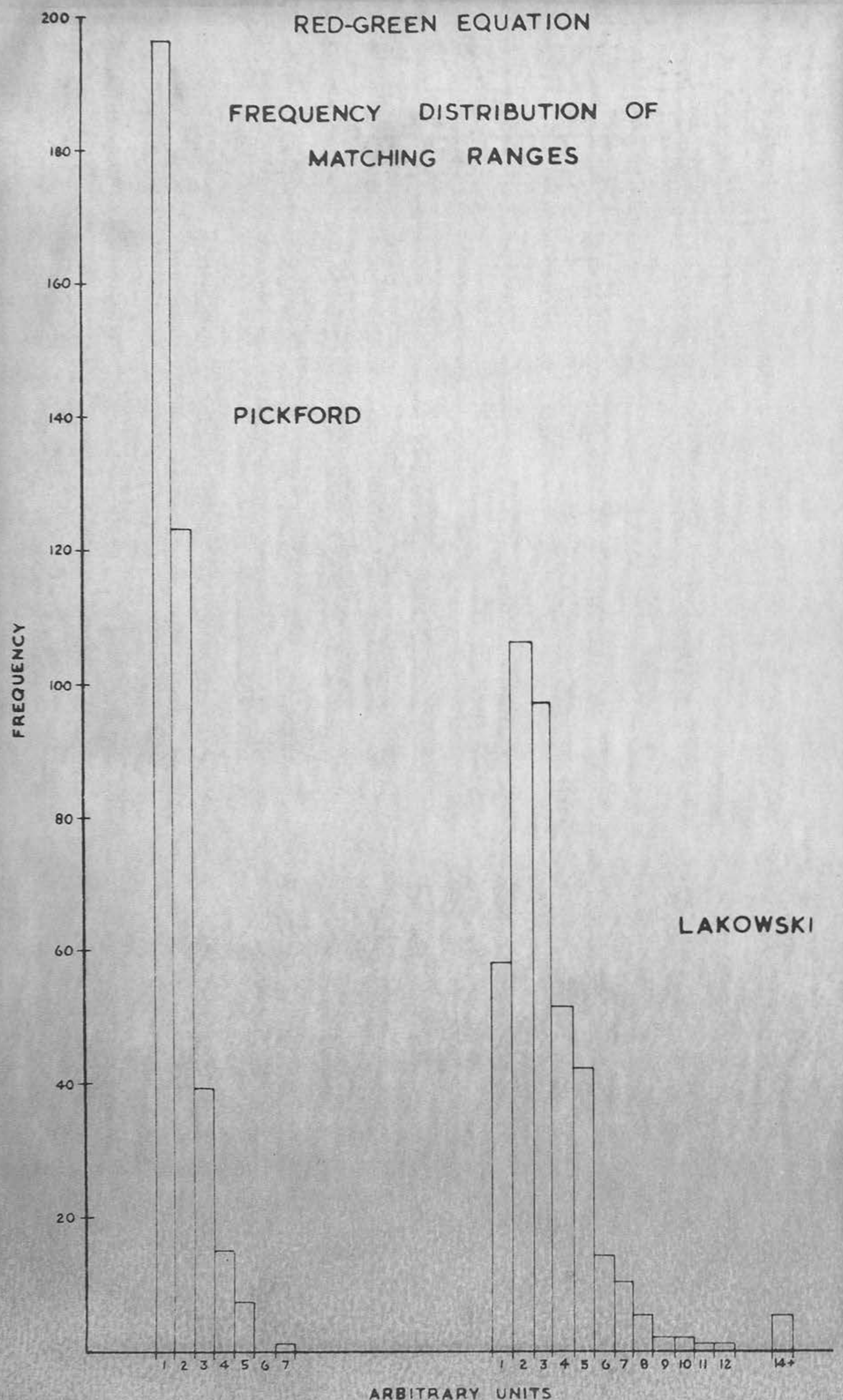
FREQUENCY DISTRIBUTION OF MID-MATCHING POINTS

performance on these two equations which is true irrespective of their differing ranges. There are differences in the incidence of those with large matching ranges in both equations. For example the matching ranges of more than half the subjects over 65 cover the whole of the violet-blue/green equation while there are only a few such subjects in the same age group for the red-green equation.

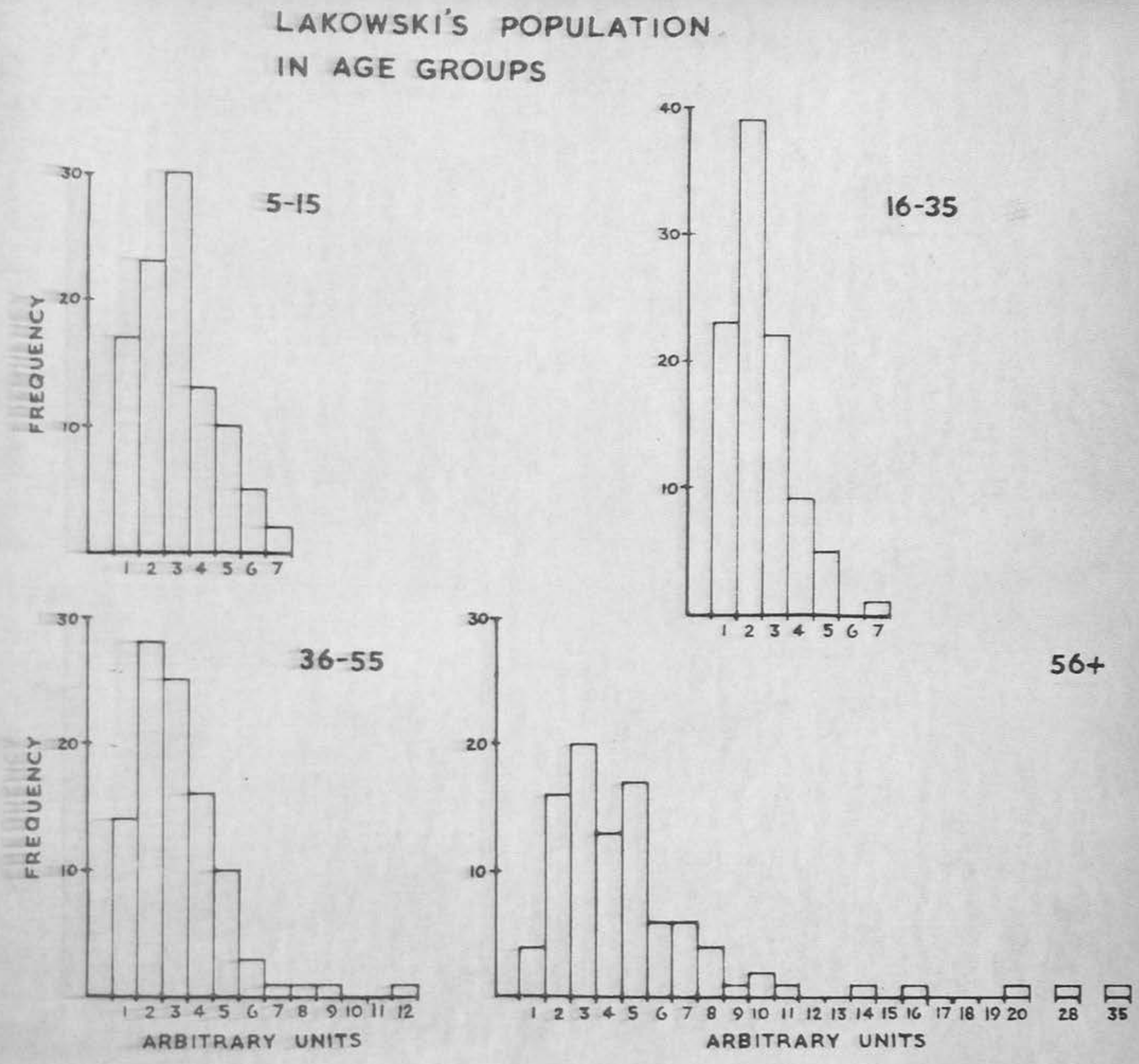
A similar comparison of the matching ranges measured in arbitrary units can be made between the yellow-blue and red-green equations. Again corrections are made for the different ranges these cover in the colour space. And again the conclusion is unavoidable that when such corrections have been taken into account, there are still more people with relatively large matching ranges in the yellow-blue equation than there are in the red-green, but not so many as were found in the violet-blue/green equation.

II. Frequency distributions - Red-green equation. Frequency distribution diagrams are given for mid-matching points and matching ranges. The first diagram gives the total frequency distribution for both Pickford and Lakowski populations. As the populations are nearly equal in size, no change in percentage presentation of frequencies was necessary. The peaks for both distributions coincide, but the spread of mid-matching values differ and the most frequent score is smaller in the Lakowski population than it is in Pickford's. A statistical analysis of Kurtosis for these two distributions was made. The results show that Lakowski's population is not significantly different from the normal frequency distribution, whereas the distribution of Pickford's population is significantly different (at 0.01 level) from the normal. In fact it is a leptokurtic distribution.


LAKOWSKI'S POPULATION
IN AGE GROUPS
5-15
16-35
36-55
56+
FREQUENCY
ARBITRARY UNITS
RED-GREEN EQUATION
FREQUENCY DISTRIBUTION OF MID-MATCHING POINTS



RED-GREEN EQUATION
FREQUENCY DISTRIBUTION OF MATCHING RANGES
PICKFORD
LAKOWSKI
FREQUENCY
200
180
160
140
120
100
80
60
40
20
1 2 3 4 5 6 7
1 2 3 4 5 6 7 8 9 10 11 12 14+
ARBITRARY UNITS



RED-GREEN EQUATION

FREQUENCY DISTRIBUTION OF MATCHING RANGES

Then there are histograms for the distribution of the age population alone. Instead of showing results for each age group, the groups have been combined to form larger units each containing approximately 100 subjects. Thus there are only four age groups - 5 - 15, 16 - 35, 36 - 55 and lastly 56 plus. The spread of mid-matching points for Pickford's population is from 15 to 25 anomaloscope units, and only the two histograms for the 5 to 35 population are similar to this. However, it is only in the second age group that the most frequent score (19) coincides with Pickford's frequency score which is 18. The other two age groups show a wider spread of mid-matching point values, though the 36 - 55 age group still has 19 as the most frequent score.

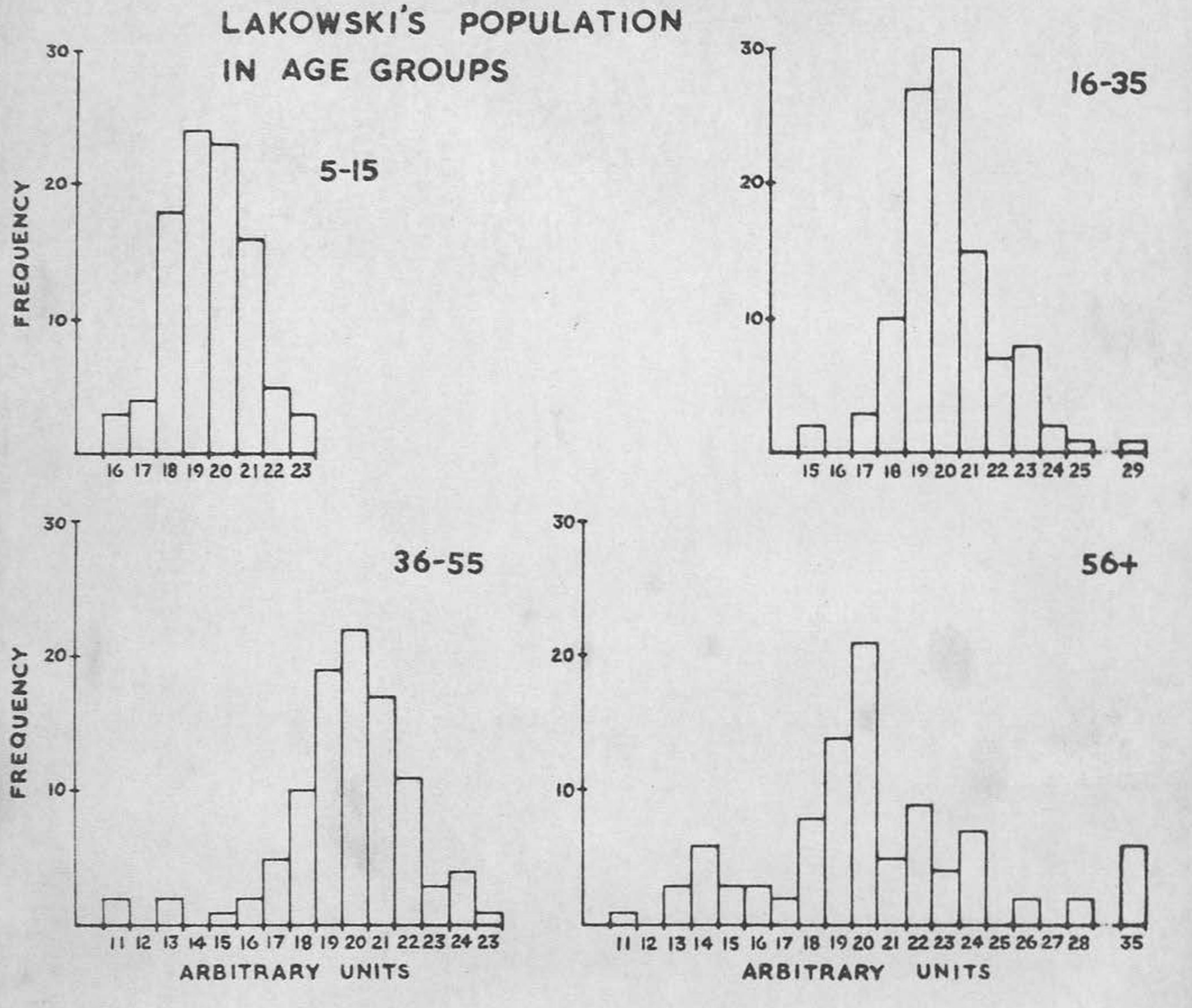
It should be stressed that in both populations all major red-green defectives have been excluded.

The next two diagrams refer to the frequency distributions of the matching ranges. One histogram gives the distributions for both populations and the other refers to the sub-groups of the age and colour vision population.

The essential point of the two comparisons is that there are three times as many subjects with matching ranges of one unit in Pickford's population than there are in Lakowski's, and in consequence, his distribution is in the form of a J curve, whereas Lakowski's is skewed.

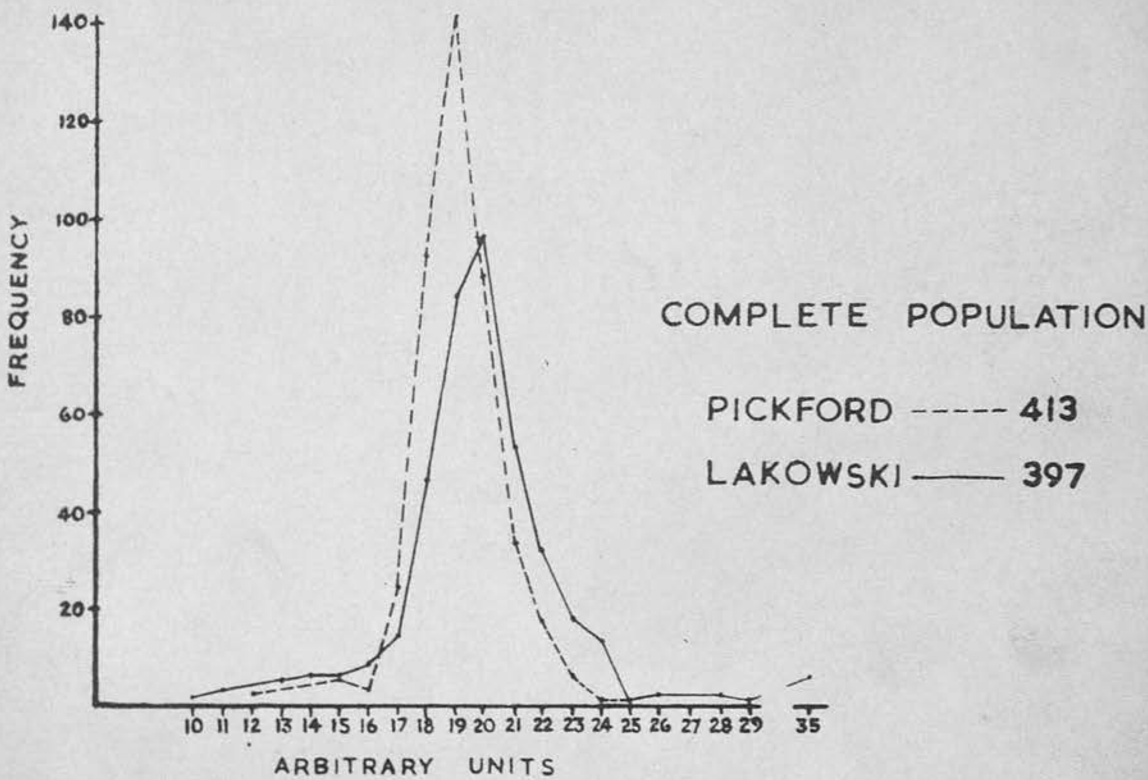
In the diagram showing the distribution of matching ranges for the four sub-groups, only the 16 - 35 age group in any way resembles Pickford's distribution.

Yellow-blue equation. Again two frequency distribution diagrams are presented for the mid-matching points for this equation. One deals with the



YELLOW-BLUE EQUATION

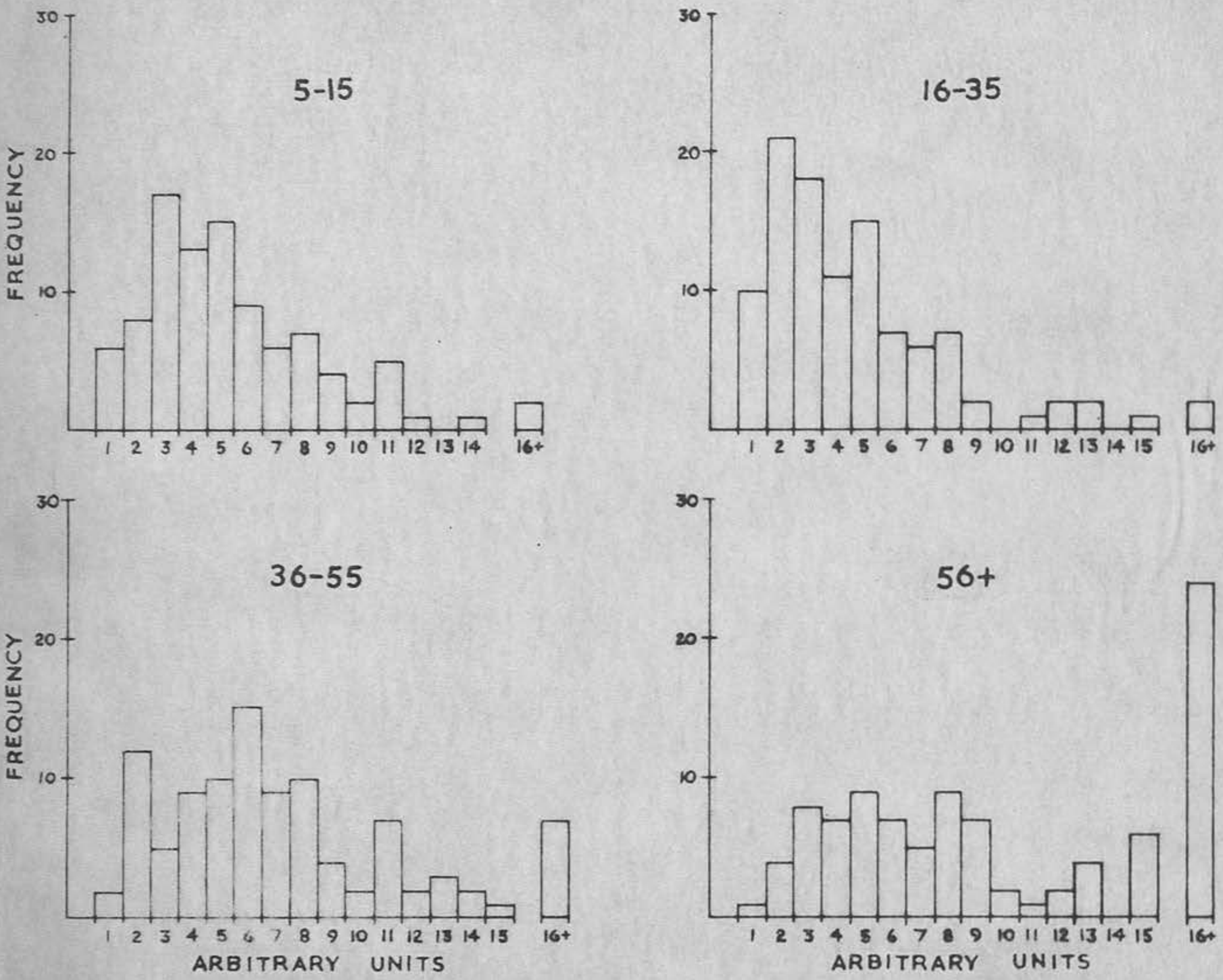
FREQUENCY DISTRIBUTION OF MID-MATCHING POINTS



YELLOW-BLUE EQUATION

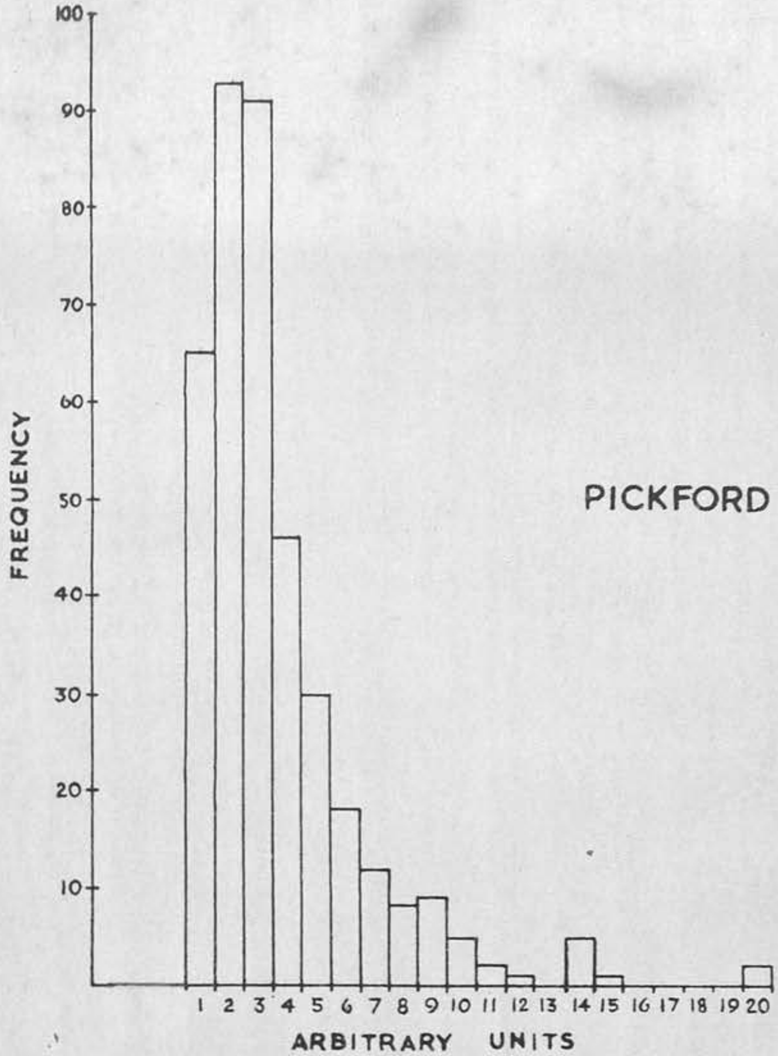
FREQUENCY DISTRIBUTION OF MID-MATCHING POINTS

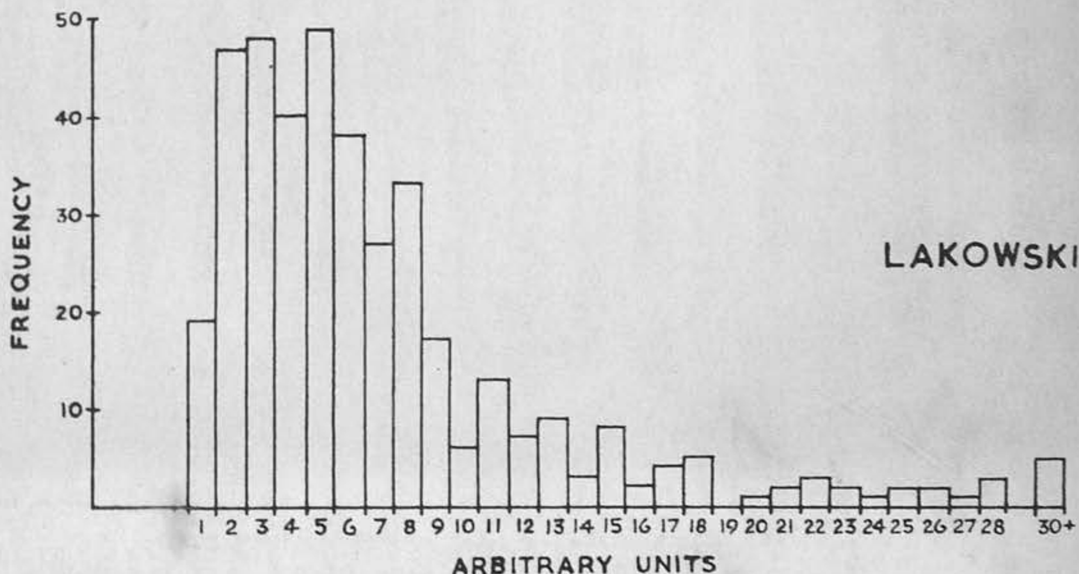
LAKOWSKI'S POPULATION IN AGE GROUPS



YELLOW-BLUE EQUATION

FREQUENCY DISTRIBUTION OF MATCHING RANGES





YELLOW-BLUE EQUATION

FREQUENCY DISTRIBUTION OF MATCHING RANGES

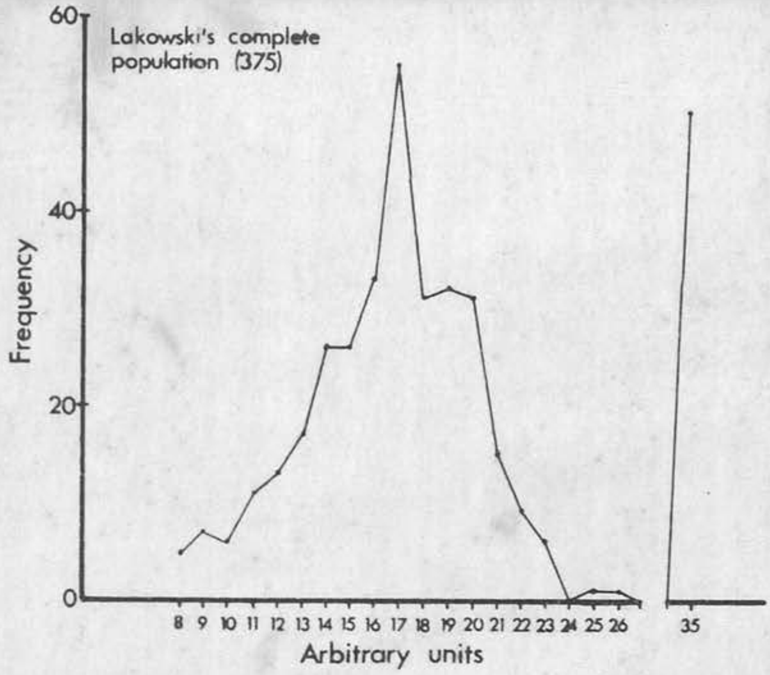
combined populations and the other with the four age sub-sections. Comparing the frequency polygons for the separate populations we find that the most frequent score is 20 in Lakowski's population and 19 in Pickford's. Again there are differences in the spread of scores and the analysis for kurtosis again indicates that the age population is a normal distribution whereas Pickford's is leptokurtic, the differences being significant at the 0.01 level.

In terms of the spread of scores, only the histogram for the 5 to 35 age group resembles Pickford's scores - all the other sub-groups have larger variations in the distribution of mid-matching points. All subjects were included in this analysis. Note that the spread of scores for the 56+ age group ranges from 11 to 28 anomaloscope units.

The frequency distribution of the matching ranges for Pickford's and Lakowski's populations are both decidedly skewed but the difference between the two is that Pickford's population is more positively skewed than Lakowski's. There are also differences in the most modal frequencies, which are 3 and 4 units in the Pickford population and at 2 to 6 anomaloscope units in Lakowski's. Again, only the 16 - 35 age group approximates to Pickford's frequency distribution of scores for this variable. Note that if we include all the subjects with 16+ range in the last group, the trend of the distribution is almost reversed - that is, it tends to become negatively skewed.

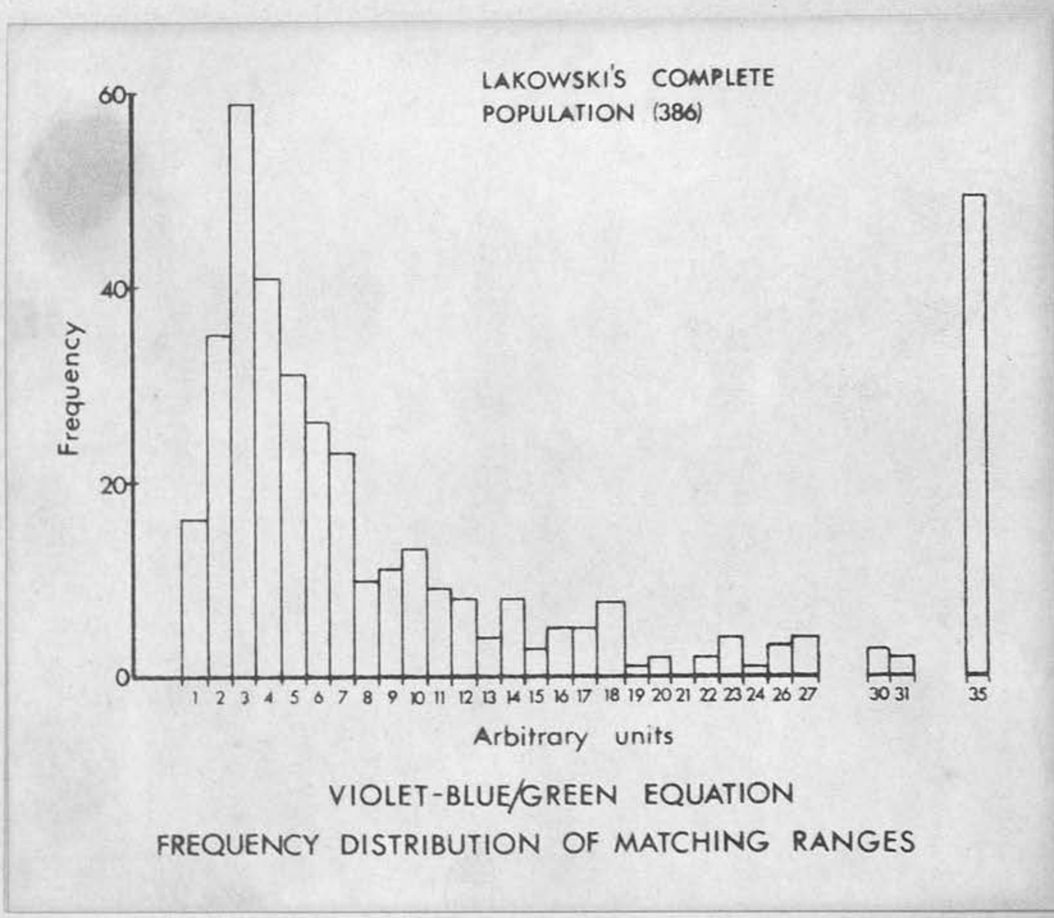
Violet-blue/green equation. Here no comparison is available, as this equation has not been studied by Pickford. Four diagrams are again presented, two for the mid-matching points, and two for matching ranges.

The frequency polygon for the total age population resembles the two previous ones, except that the normal distribution curves are now less smooth.

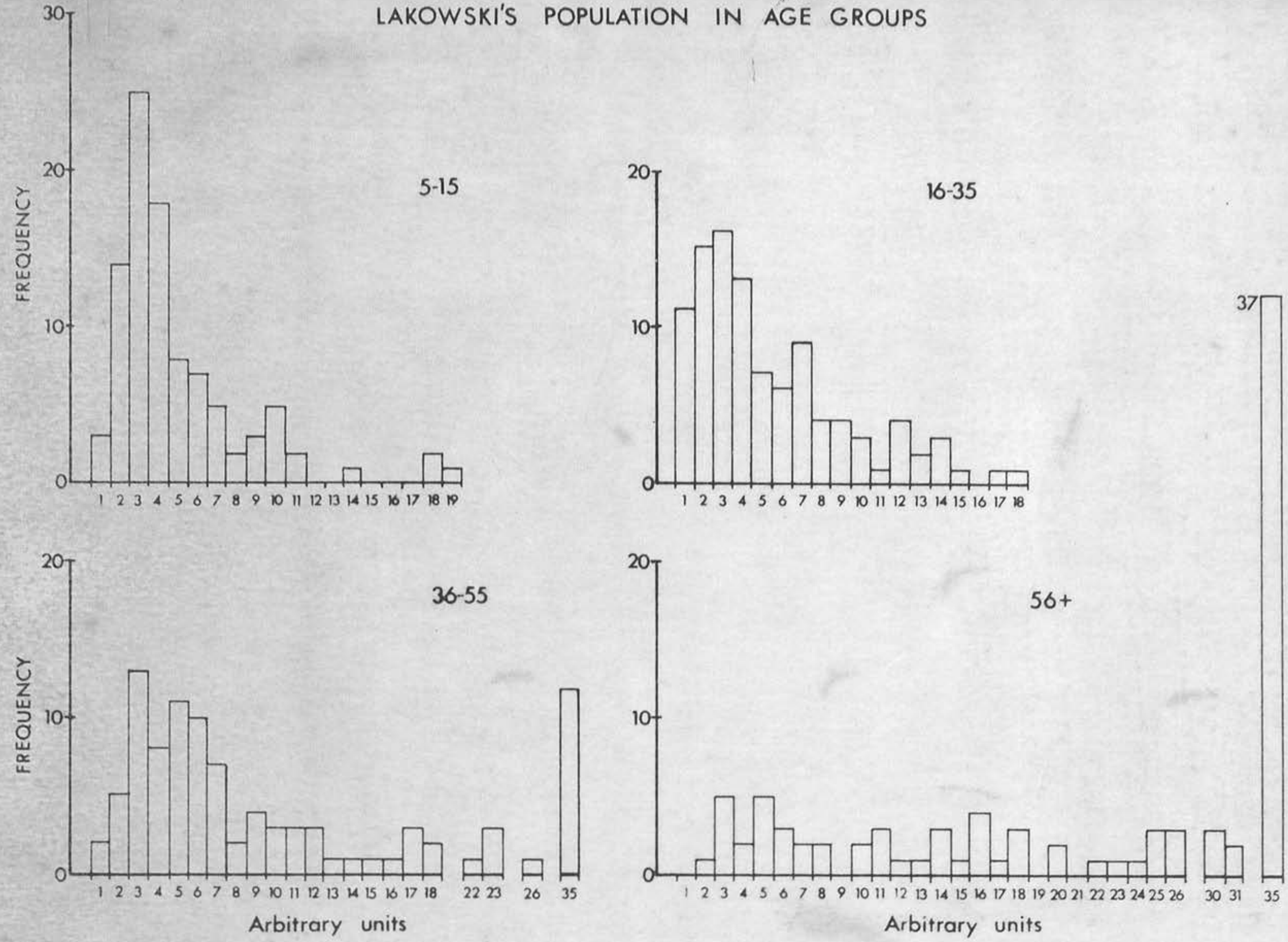


VIOLET-BLUE/GREEN EQUATION

FREQUENCY DISTRIBUTION OF MID-MATCHING POINTS


LAKOWSKI'S COMPLETE
POPULATION (386)
Frequency
0
20
40
60
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 26 27 30 31 35
Arbitrary units
VIOLET-BLUE/GREEN EQUATION
FREQUENCY DISTRIBUTION OF MATCHING RANGES

LAKOWSKI'S POPULATION IN AGE GROUPS



VIOLET-BLUE/GREEN EQUATION

FREQUENCY DISTRIBUTION OF MATCHING RANGES

In this equation there are almost as many subjects who match throughout the whole equation as there are people with the modal mid-matching point at 17. The frequency for the first group has been placed at the end of the scale (i.e. 35) since it is unacceptable to give them a modal score.

The frequency polygons for the four sub-groups show shifts from one sub-group to the next in the modal frequencies in this equation. Note that there is a tendency towards a platokurtic distribution as we reach the older age groups. The frequency of those accepting matches equivalent to the range of the entire equation is shown for the two last sub-groups, where they appear at the 35 point position on the graphs, (the frequency for the 36 to 55 group is about 12 while it is as much as 38 for the last age groups).

In the diagram dealing with matching ranges there is a positively skewed distribution of scores for the total population. This distinct pattern of skewness is present in the two age groups from 5 to 35 years of age, a vestige still remains in the 36 - 55 group but disappears in the last group, where the frequency distribution is almost equally spread over the entire range from 2 to 31 arbitrary units.

III. <u>Statistical analysis</u> - The results from testing subjects with the anomaloscope were expressed in two measures for each colour equation. Statistical analysis of these measures is complicated by the fact that the matching ranges do not follow the normal distribution curve, and therefore it is not possible to use the usual techniques of statistical presentation. Instead non-parametic methods have to be employed for analysing matching ranges. The mean and standard deviation technique has, however, been used in the analysis of mid-matching points.

AGE	RED-GREEN			YELLOW-BLUE			VIOLET-B/GREEN		
	N	Mean	S.D	N	Mean	S.D	N	Mean	S.D
5-10	47	18.1	1.5	47	19.3	1.7	47	16.1	2.6
11-15	47	18.7	1.6	47	19.1	1.9	46	17.8	2.9
16-25	50	19.4	1.6	49	20.0	2.3	47	16.8	3.5
26-35	50	19.3	1.1	50	20.0	1.7	50	17.1	3.1
36-45	50	19.1	2.3	49	19.5	2.4	44	15.9	3.1
46-55	50	19.2	1.7	50	19.4	2.2	45	14.5	3.2
56-65	44	18.8	2.4	44	19.4	2.7	32	15.2	2.9
66+	46	19.1	2.6	41	18.7	3.3	21	15.3	4.0

TABLE 10

Means and Standard Deviations of the Mid-Matching Points.
N indicates number of subjects in the original eight age groups.

EQUATION	PARAMETER	N	5-15	N	16-35	N	36-55	N	56+	N	Total	N	Pickford
R-G	Mean	94	18.4	100	19.4	100	19.1	90	19.0	384	19.0	394	19.1
	S.D		1.6		1.4		2.0		2.4		1.9		1.3
Y-B	Mean	94	19.2	99	20.0	99	19.5	85	19.1	377	19.5	436	19.6
	S.D		1.8		2.0		2.3		3.1		2.4		1.7
V-B/G	Mean	93	17.0	97	17.0	89	15.2	53	15.3	332	16.2		---
	S.D		2.9		3.3		3.3		4.6		3.5		---

TABLE 11

Means and Standard Deviations of the Mid-Matching Points.
Results of the four compound age groups, the total sample and Pickford's sample.

Originally the data was divided into eight groups, and later to increase the number of subjects in the samples, it was combined into four age groups. Table No. 10 records the means and standard deviations of the mid-matching points for the three colour equations, on the basis of the original eight age groups, and Table No. 11 for the four sub-groups and Pickford's sample.

Each of the eight age groups was tested to find if it showed a statistically significant difference from the 16 - 36 group. The choice of such a 'reference group' was thought advisable, because it most closely resembled the age scatter in Pickford's population and because the ageing effect was thought to be unlikely. It will be seen from Table No. 12 that the red-green and yellow-blue equations, the student 't' test indicates significant differences only between the very young age groups and the 16 - 35 age group. The means of all the other mid-matching points remain unchanged. In the violet-blue/green equation the means for the mid-matching points move toward the violet part of the equations with each successive age group.

A more detailed investigation of the significance of the differences in mid-matching points was conducted with the data in Table No. 11. Each of the four groups was tested against one another both for differences in mean and standard deviation with results as shown in Table No. 13.

Table No. 14 shows the same analysis calculated by means of the F ratio based on variance. Essentially the results are similar to those conducted with the smaller age groups but some refinements can be noted. For example, in the red-green equation there are significant differences in dispersion between the 16 - 36 group and the two older ones, although the means are not significantly

EQUATION	5 - 10	11 - 15	36 - 45	46 - 55	56 - 65	66+
RED-GREEN	.01	.05	NS	NS	NS	NS
YELLOW-BLUE	.05	.05	NS	NS	NS	.05
VIOLET-BLUE/ GREEN	NS	NS	NS	.01	.01	NS

TABLE 12

Significance levels of differences between the Means of the Mid-matching Points of the age groups shown and the combined 16 - 25 and 26 - 35 age groups as determined by Student 't'.

AGE GROUP		5 - 15		16 - 35		36 - 55		56+	
		M	SD	M	SD	M	SD	M	SD
RG	16-35	.01	NS						
	36-55	.01	.05	NS	.01				
	56+	NS	.01	NS	.01	NS	NS		
	P'ford	.01	.05	.05	NS	NS	.01	NS	.01
YB	16-35	.01	NS						
	36-55	NS	.05	NS	NS				
	56+	NS	.01	.05	.01	NS	.01		
	P'ford	.05	NS	NS	NS	NS	.01	NS	.01
V-B/G	16 - 35	NS	NS						
	36-55	.01	NS	.01	NS				
	56+	.05	.01	.05	.05	NS	.05		

TABLE 13

Significance levels of differences between the Means and Standard Deviations of the Mid-Matching Points of the age groups shown.

	AGE GROUP	5 - 15	16 - 35	36 - 55	56+
RG	16-35	.01			
	36-55	.01	NS		
	56+	NS	NS	NS	
	P'ford	.01	.05	NS	NS
YB	16-35	.01			
	36-55	NS	NS		
	56+	NS	.05	NS	
	P'ford	.05	NS	NS	.05
V-B/G	16-35	NS			
	36-55	.01	.01		
	56+	.01	.01	NS	

TABLE 14

Analysis of Variance of Mid-Matching Points. Significance determined by means of the F test.

different. In yellow-blue, the 16 - 35 group differs significantly from the oldest and youngest groups in their means. Finally, in violet-blue/green there are differences throughout except between the two younger and the two older groups. The last finding may be affected by the small number of results in the 56+ group.

The mid-matching point can be summarised as follows :-

Red-green. Significant changes in the mean value occur for the younger subjects but then it settles around a fixed value after 36. The minimum dispersion is in the 16-35 age group, which is therefore the most stable group. Pickford's sample which as we saw earlier shows a rather different distribution profile is similar to the groups over 36 years of age.

Yellow-blue. There is greater variation in the results for this equation, and so any probable differences between means are largely invalid. Pickford's sample is not significantly different from the results of subjects over 16 years.

Violet-blue/green. Despite the somewhat larger standard deviations, most of the means for the four age-groups are significantly different, so that this equation would appear to be most affected by age.

An analysis of differences in matching ranges with age can be made in two ways, one by considering the frequency distributions previously presented in section 3:5(b) II, and the other using a special statistical technique devised by Festinger (1943) to assess the significance of differences between progressively skewed distributions.

Firstly, referring to diagram p. 237-8 for the red-green equation, it

	AGE GROUP	5 - 15	16 - 35	36 - 55	56+
RG	16-35	.01			
	36-55	NS	.01		
	56+	.01	.01	.01	
	Pickford	.01	.01	.01	.01
YB	16-35	NS			
	36-55	.01	.01		
	56+	.01	.01	.01	
	Pickford	.01	.01	.01	.01
V-B/G	16-35	NS			
	36-55	.01	.01		
	56+	.01	.01	.01	

TABLE 15

Significance of differences between the skewed distribution of Matching Ranges using Festinger's method.

is apparent that the 16 - 35 age group has the smallest spread. The modes of the two middle age groups fall around the position of the second arbitrary unit, whilst the other two groups have theirs in the third unit. According to the calculations in Table No. 15 these four distributions are significantly different except for 5 - 15 compared with 36 - 55 age groups.

The yellow-blue results in diagram p.238-9 show very little difference between 5 - 15 and 16 - 35, which is confirmed by a lack of significant difference. The next two age groups are much more evenly distributed and the 56+ group has many subjects with ranges of sixteen or more units. Both groups are significantly different from all others.

Thirdly, the violet-blue/green results in diagram p.240 exhibit a gradual deterioration from the youngest age group to the oldest with a very pronounced loss of discrimination in the 56+ group as shown by the large number of subjects with a range of over 35 units. Again the four groups are significantly different from one another, except the first two in relation to each other.

As was seen with the mid-matching points the violet-blue/green equation seemed to be most affected by age. But both the red-green and yellow-blue equations were also affected by age once the subjects were over 35 years.

An analysis according to sex was also carried out. Significant differences between the standard deviations but not the means, of the mid-matching points for the total populations of males and females were found in the red-green and yellow-blue equations (in each case the females having the larger dispersion). However, an analysis of variance on mid-matching points yielded no significant differences. On the other hand, an analysis of variance of matching ranges

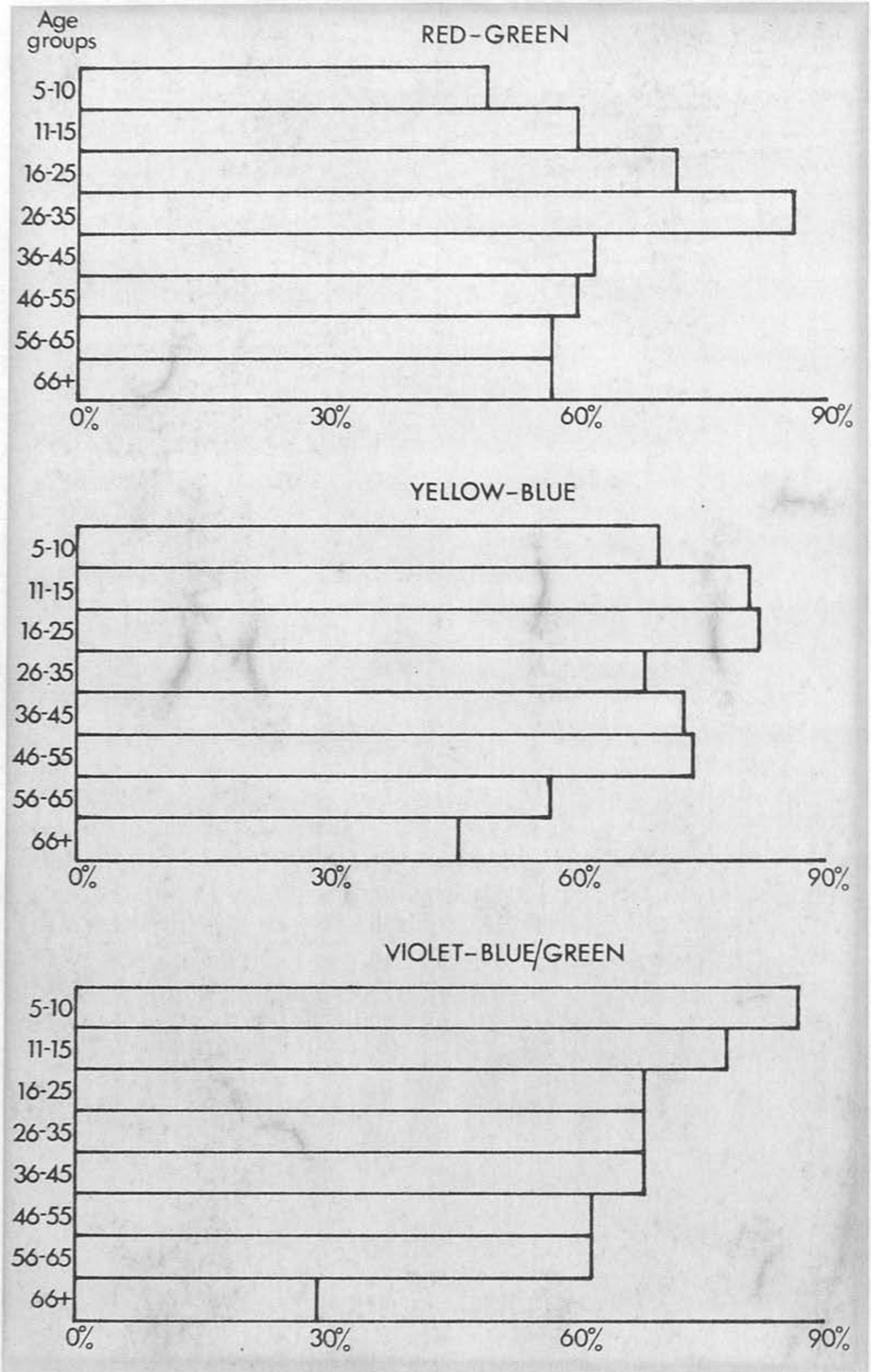

Age groups
RED-GREEN
5-10
11-15
16-25
26-35
36-45
46-55
56-65
66+
0%
30%
60%
90%
YELLOW-BLUE
5-10
11-15
16-25
26-35
36-45
46-55
56-65
66+
0%
30%
60%
90%
VIOLET-BLUE/GREEN
5-10
11-15
16-25
26-35
36-45
46-55
56-65
66+
0%
30%
60%
90%

DIAGRAM 27

showed significant differences between males and females in the yellow-blue and violet-blue/green equations.

Within individual age groups, significant differences in the means of the mid-matching points were found in red-green for the 11 - 15 group, and in yellow-blue for the 16 - 35 group. This might be due to the differences in the ages of males and females at the onset of puberty - Sorsby (1961).

The female population certainly seems to show more varied results on the anomaloscope.

A similar picture of the effect of age upon the performance of subjects tested with the anomaloscope can be obtained in the following way.

Diagram No. 27 gives percentages for subjects in each of the eight age groups whose mid-matching points are within $\pm$ S. D. of the mean of the 16 - 35 age group, where colour discrimination is thought to be most sensitive. The red-green equation shows remarkably little age difference except in the 16 - 35 group, but the other two equations show gradual changes with age where fewer and fewer subjects have their mid-matching points within the limits set by the 16 - 35 group.

4. SUMMARY AND CONCLUSIONS

The historical account gave us an insight into the difficulties that faced early research workers in their studies of the effect of ageing upon colour discrimination. The main obstacles were the lack of a proper definition of the concept of colour vision; failure to understand the principles underlying the construction of colour vision tests, and the tendency to approach such research with very narrow and rigid a priori ideas of what could be expected in an age population.

The results of this thesis show that contrary to what has been said before, even pseudo-isochromatic tests can give some indication of what happens among older subjects. Some test items and pseudo-isochromatic plates were found to be more sensitive in detecting age variations than others. Individual plates in the Dvorine and Ishihara tests give a very good indication of changes in colour perception. However, though this is true for particular items, these tests on the whole are rather crude measures of the smaller changes occurring from one age group to another. It is questionable whether sex differences in performance are valid differences since the finding for the Ishihara contradict these of the Dvorine. The results will differ according to what a given test is testing. Discrepancies between these two tests are mainly due to the fact that in the Ishihara the yellow-blue element is less in evidence, and it is this discrimination (more typical in the Dvorine) that first shows signs of ageing.

Judging by the results obtained on these tests, this happens earlier and more often among women than men. Though the tritan plates conform with this general trend not all show a definite age variable such as was found in the

F. 5 plate.

Because of its inherent design and because it measures colour in three equations the Pickford anomaloscope is a very fine test. The main axes of the colour triangle offer a very sensitive method of detecting both the minute and gross changes in colour discrimination that occur in an age population. Anomaloscope testing has shown that there are important variations in the ratios for the three equations among subjects up to 15 years of age. Thus not only is their colour discrimination less accurate than those in the 20 to 30 age group, but they also differ in terms of the most frequent ratios accepted.

The age effect was evident from the fact that matching ranges increased significantly in all equations from age group to age group after 25. The greatest loss in discrimination was observed for the violet-blue/green equation. Another important contribution resulting from the anomaloscope studies was the emergence of a group of people, who in terms of colour matching, resembled Extreme Deuteranomalous or Protanomalous subjects. Their matching ranges though starting within the limits of the normal distribution for mid-matching points, went beyond these in both directions and this might occur in any of the three equations. It is important to stress that in contrast to the work of Boles-Carenini which showed a bias towards the green end of the spectrum in the red-green ratios of the older subjects, it was found that this bias extended in both directions - i. e. both towards the red and green. This is an important finding, since if the shifts of ratios were observed in the green part of the spectrum, the classical explanation that colour changes in aged subjects are the outcome of pre-retinal changes, could be accepted as the only physiological cause for this deterioration. The fact that there are as many subjects with shifts to the red-ends of the equations forces us to look for additional explanations for the causes of ageing. The old dictum that all changes are 'normal senile' changes can no longer be accepted.

An analysis into sex variations showed that for the total population, there was a significant difference between the standard deviations of the mid-matching points - females showing a larger dispersion on both the red-green and yellow-blue equations. However, within individual age groups, significant differences were found in the means of the mid-matching points in the red-green equation for the 11 - 15 age group; and in the yellow-blue equation for the 16 - 35 age group.

Photometric and colorimetric analysis revealed that only the use of these methods of analysis will really indicate for a given test what variables are being tested, and where these lie in the colour space. Comparisons to find which of a series of tests is the 'best' are not sufficient if judged by performance alone, for such comparisons are only admissable when adequate information about these tests is available. Only in the light of this data is it possible to correlate the results of the various tests.

It is safe to assume that no two tests measure the same variable in the same manner. There is always some difference, and even if this is small it might affect any inferences we make about a particular subject's colour discrimination.

The photographic method outlined here offers a quick means of checking on one of the important variables in colour vision - namely the relative luminosity of the individual colours in a given test.

Has any additional evidence come to light since 1958 to verify the findings of this thesis ?

Crawford's study (1959) seems to confirm the interesting deviations that were observed among the very young. The object of his study was to determine by direct experiment the degree of variation which could be tolerated in an illuminant, intended to conform to a given type (say daylight). The criterion used was just noticeable differences from the standard, judged by memory alone. The instrument used was a double monochrometer, where spectral stimuli are used.

He grouped his results according to age, employing 4 age groups each with 10 observers, from 16 to 21, from 22 to 28, from 29 to 38, and from 45 to 64, and the characteristic shape of the curve was found to be almost identical for all age groups. Age differences, though not large, were found. The results of the oldest subjects showed a reduction in colour vision, and losses of sensitivity to the violet end of the spectrum were indicated by higher tolerance in comparison with the previous age group. This could be predicted from the known yellowing of the lens of the eye with increased age. The results for the youngest group did not follow the general pattern but Crawford did not offer any explanation for this in his account.

There is further evidence for the different ratios accepted by the younger subjects to be found in Lakowski's (1962) study on 200 printers' apprentices between the ages of 13 and 16. This study was carried out using the Pickford-Nicolson anomaloscope and a battery of pseudo-isochromatic plates similar to those used in this thesis. The apprentices' results were compared with those of a control group consisting of 55 subjects whose mean age was 30 years. Again it was found that important differences between these two groups were found not only for the matching ranges, but also in the most frequent mid-matching points of the two groups. The means of the mid-matching points of the younger subjects (in this case the apprentices) were significantly (at 0.01 level) shifted towards the red in the red-green equation, and to the yellow in the yellow-blue equation.

The results of both groups for the pseudo-isochromatic tests were very similar to the findings of the study presented for this thesis.

Evidence from two other studies of age populations again confirm the general trends outlined here.

Evidence from two other studies of age populations again confirm the general trends outlined here.

The study of P. Speciale Picciche and F. Bozzoni (1961) using the CAT on 46 subjects arranged in 3 age groups (20 to 25, 40 to 50 and over 60 years of age), showed that subjects over 60 years made more mistakes in the CAT than younger subjects, thus, confirming the findings of Gilbert and Quellette.

In 1962 Verriest and his collaborators described the results of a study of colour and age differences using the Farnsworth Munsell 100-Hue test. In this experiment 480 subjects between the ages of 10 and 65 were tested. All subjects with congenital deficiencies as tested by the A.O. or the H.R.R. tests were excluded, and the results grouped according to colour, sex and age, (subjects were divided into eleven half decade groups).

Mean differences for each decade in terms of sex were analysed first and it was found that differences were only reliable in the 15 to 24 year decade where the women scored significantly less than the men. When sex was disregarded the eleven mean total scores of the age groups were significant, and this was especially true of the 20 to 24 years group which was the lowest of them all.

The mean total increased more or less regularly on either side of this group. Verriest thought that, psychological factors might be responsible for the high scores in the very young subjects but the differences could be attributed to a true physiological heterogeneity.

Verriest's general conclusion s were that hue discrimination as tested by the 100-Hue test is best in young adults and definitely worse in the older age groups who appeared to be especially weak in the blue-green and reds just like congenital Tritanomalous subjects. Nothing could be said about the comparison of normal male and female trichromats.

Conclusions. The purpose of the present study has been to find out whether there were any changes in colour vision due to age. Over 500 subjects were tested on a battery of pseudo-isochromatic plates and on the Pickford anomaloscope. The age range was from 5 to 90 years.

From the data collected it appears that there are distinct phases in the development of colour discrimination. Colours are perceived most readily and most accurately in the younger age groups. On the anomaloscope the best age group for the red-green equation is 26 - 35; for yellow-blue it is the 16 - 25; and for violet-blue/green it is 5 to 15. On the pseudo-isochromatic plates those in the 20 to 30 age group make the least number of misreadings.

Lack of this accute colour perception in other age groups seems to have causes which are distinct for the younger and for the older groups. Most research workers emphasize that the psychological maturation of the perceptual process is the main cause of the relatively poor colour discrimination of the young. Evidence from the anomaloscope showing how their modal mid-matching points differ from that of the adults, points to the interesting possibility that their basic visual mechanism - or its optical components - must be different, thus adding a physiological factor to the psychological one.

In the older subjects, the changes detected on the anomaloscope have to be related to the various changes in the optical system, changes in the properties of the lens and in the pigmentation of the macular region. Essentially all such losses may be symptoms of senescence, but many of the matching patterns of the older subjects are 'too bizarre' to be explained in these terms alone. A better explanation would be that pathological changes are more common in older people than is admitted at present.

Results of the work of Francois and Verriest (1957) and Zanen (1959) on clinical patients with ocular diseases arising from systemic disturbances of constitutional, metabolic, endocrine or vascular origin, show that these give rise to acquired deficiencies (or dyschromatopsias) in colour discrimination, which are different from genetically determined defects. The results of many older subjects for this research exhibit responses on the various colour vision tests (especially on the anomaloscope) that are similar to those described as specific clinical cases.

Perhaps it would not be an exaggeration to say that the usefulness of these findings is not confined to the industrial field alone where their relevance is obvious but that in the future, it is in the field of diagnosis that they may prove to be most useful.

BIBLIOGRAPHY.

ABNEY & FESTING. Proc. Roy.Soc., B,50,369, 1891.

ADAMS, E.Q. X-Z planes in the 1931 I.C.I. system of colorimetry. J.Optical Soc.Amer. (1942) 32, p.168.

ADAMS, J.M. Some Results Obtained with the I.S.C.C., Color-Aptitude Test. Patra (1956).

AHLENSTIEL, Archiv.f.Opthal. (1943). (As quoted by K. Janouskova,1955).

AUBERT, Physiologie der Netzhaut (Breslau) 1865.

BELCHER, S.J., K.W. GREENSHIELDS & W.D. WRIGHT. Colour Vision Survey Using the Ishihara, Dvorine, Bostrom & Kugelberg, Bostrom and American Optical Hardy-Rand-Rittler Tests.

BINET, S. Perception d'enfants. Rev.Philos.(1890) pp 30-583.

BOICE, M.L., M.A. TINKER & D.G. PATERSON, Colour Vision and Age. Amer.J. Psychol. (1948)61, pp 520-526.

BOLES-CARENINI, B. Del Comportamento Del Senso Cromatico In Relazione All'eta. Ann.Ottal., 80 (1954) pp 451-458.

BROWN, R. An Investigation into the Colour Vision of School Children. Br.J. Ed.Psych. (1950).

BURNHAM, R.W., R.M. HANES & C.JAMES BARTLESON. Colour : A Guide to Basic Facts and Concepts. J. Wiley & Sons, Inc., New York : London. 1963.

CHAPANIS, A. A comparative study of five tests of color vision. J.Opt.Soc. Amer., (1948) 38. pp 626-649.

CHAPANIS, A. Relationships between Age, Visual Acuity and Color Vision. Human Biol. (1950).

CHASE, W.P. Color vision in infants. J.Exper.Psychol. (1937) pp 203-222.

CLARK, F.J.J. Extra-Foveal Vision and the Additivity Concept. Die Farbe 11 (1962) Nr.1/6.

COLLINS, G.L. Color Blindness : Its Relation to Other Ocualr Conditions, and the Bearing on Public Health of Tests for Color Sense Acuity. Publ.Hlth.Bull. No.92. Washington: Gov't Printing Office (1918) pp 29.

COLLINS, M. Colour-blindness with a Comparison of Various Methods of Testing Colour-Blindness. London : Trubner, (1925).

COLLINS, M. Tests in Common Use for the Diagnosis of Colour Defect. Ann.Rep. Brit.Ass. (1937) pp 208-226.

THE COLOUR GROUP, Committee of, Report on Defective Colour Vision in Industry. The Physical Society, London, 1946.

COOK, W.M. Ability of children in color discrimination. Child Devel.(1931) 2, pp 303-320.

CRAWFORD, A. The Dvorine Pseudo-Isochromatic Plates. Brit.J.Psychol.(1955).

CRAWFORD, B.H. Measurement of Color Rendering Tolerances. J.Opt.Soc.Amer. 49(12), (1959).

CRONSTEDT, L. Nagra iakttagelser angaende behovet av glaskorrektion vid fargsinnesundersokningar. Svenska Lakartidn. (1937), 34, pp 145-150.

DAVIES, A., & W. PIERCE. The Selection of Colour Workers. London : Pitman (1934).

DE KERF, J.L.F. Accuracy of Tristimulus Computations. J.Opt.Soc.Amer. (1958) 48, No.5.

DIMMICK, F.L. Specifications and Calibration of the 1953 Edition of the Inter-Society Color Council Color Aptitude Test. J.Opt.Soc.Amer.(1956) 46, 6.

DOW. Phil.Mag,12,120,1906.

DVORINE,I. Pseudo-Isochromatic Plates. Sec.Ed. Scientific Publishing Co., 2328 Eutaw Place, Baltimore 17, M.D. (1955).

FARNSWORTH, D. The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. J.Opt.Soc.Amer. (1943) 33, 568.

FARNSWORTH,D. The Farnsworth Dichotomous Test for Colour Blindness, Panel D-15. The Psychological Corporation, 522 Fifth Avenue, New York 36. N.Y.

FARNSWORTH, D. The Farnsworth-Munsell 100-Hue Tests for the examination of Colour Discrimination. Munsell Color Company, Inc., 2441 North Calvert Street, Baltimore, Maryland.

FARNSWORTH, D. - private communication (letter dated 11th March, 1959).

FEDERATION OF SOCIETIES FOR PAINT TECHNOLOGY. Color Aptitude Test, 121 South Broad Street, Philadelphia P.A.

FERGUSON, G.E. Statitical Analysis in Psychology and Education. McGraw-Hill Book Co. Inc., New York, 1959.

FERRARA, A. Soglia visiva per il bleu nei giovani, nei vecchi e negli afachici. Ann. Ottalmol.e Clin.Oculist. (1940) 68, pp 147-154.

FESTINGER, L. A statistical test for means of samples from skew populations. Psychometrika. Vol.8, Dec.(1943).

FOSTER, H. A Comparative Study of three Tests for Color Vision. J.Appl. Psychol. 30, (1946) pp 135-143.

GARBINI, A. Evoluzione del senso chromatico nei bambini. Arch.per l'antr. (1894) 24, pp 71-88.

GIBSON, K.S. Spectrophotometry. NBS Circular 484 (1949).

GILBERT, J. Researches on the mental and physical development of school children. Stud.Yale Psycho.Lab., (1894) 2, pp 40-100.

GILBERT, J.G. Age Changes and Colour Matching, J.Gerontology B.2 (1957) Vol.12.

GRANVILLE, W.C. Dynamic Metamerism presented at I.S.C.C Annual Meeting, New York City (1962).

GRANVILLE, W.C. Metamerism and Color Stability. I.S.C.C Newsletter No.165-166. pp 20-24.

GUILFORD, J.P. Fundamental Statistics in Psychology and Education. McGraw-Hill Book Co.,Inc. 1956.

HALSEY, R.M. & A. CHAPANIS. An Experimental Determination of some Isochromaticity Lines in Color-Deficient Vision. J.Opt.Soc.Amer. 42(lo) 1952.

HARDY, L.H., G. RAND & M.C RITTLER. The Ishihara Test as a Means of Detecting and Analyzing Defective Colour Vision. J.Gen.Psychol.(1947) 36,pp 79-106.

HARDY, L.H., G. RAND & M.C. RITTLER. AO-HRR Pseudo-Isochromatic Plates for Detecting, Classifying and Estimating the Degree of Defective Colour Vision. American Optical Co., Southbridge. (1955).

HAYES, S. Colour defects: Measurement, classification, heredity. Psycho. Bull. (1926) 23, pp 547-573.

HAZLITT, S. Children's thinking. Brit.J.Psychol. (1930) 20, pp 354-361.

HEATHCOTE, A.G.S & D.S. MACPHERSON. The Incidence of Color Vision Defects in Toronto Grade VI School Children. Can.J.Pub.Health.Vol.54, Feb (1963).

HELMHOLTZ,H.v. Handbuch der Physiologischen Optik. Leopold Voss, Hamburn und Leipzig, 1911.

HESS. Archiv.f.Augenheil (1908-9) (As quoted by K. Janouskova,1955).

HOUSTON, R.A. Vision and Colour Vision. New York: Longmans, (1932).

IVES, H.E. Studies in the Photometry in Lights of Different Colours. Phil.Mag.24, pp 70-352.

JANOUSKOVA, K. Barevne Videni a Vek.Cs.Oftal. 11 (1955) pp 37-48.

JOHNSON. Trans. Ill. Eng. Soc., 32 646 (1937).

JUDD, D.B. A Maxwell Triangle yielding uniform chromaticity scales. J. Opt.Soc.Amer.(1935) 25,24.

JUDD, D.B. Colorimetry. NBS Circular 478 (1950).

JUDD, D.B. Basic Correlates of the Visual Stimulus. Handbook of Experimental Psychology. pp 811-867.

JUDD, D.B., L. PLAZA & D FARNSWORTH. Tritanopia with abnormally Heavy Ocular Pigmentation. J. Optical Soc. Am. (1950) Vol.40, pp 833-841.

JUDD, D.B. and WYSZECKI, G. Colour in Business, Science and Industry. John Wiley and Sons Inc., New York & London, 1963.

ISHIHARA, S. Tests for Colour-Blindness. H.K. Lewis & Co., Ltd. London 9th Ed.

KALMUS, H. The Genetics of Tritanopia. Atti.Del IX Congresso Internazionale di Genetica, Caryologia, Vol.Suppl.1954.

KALMUS, H. The Familial Distribution of Congenital Tritanopia. Ann. of Human Genetics (1955) Vol.20. pp 39-56.

KELLY, K.L. Observer Differences in Color-Mixture Functions Studied by Means of a Pair of Metameric Grays. J.Res.Nat.Bur.Stand., 60 (2) (1958).

KLEEMEIER, R.W. The Relationship Between Ortho-Rater Tests of Acuity and Color Vision in a Senescent Group. J.Appl.Psychol. (1952).

KOFFKA, K. The Growth of the Mind. New York: Harcourt Brace (1924).

LAKOWSKI, R. Age and Colour Vision. The Adv. of Science.V.15 (1958) pp 231-237.

LAKOWSKI, R. Experiments in Colour Vision, Especially with Reference to Aging. Bull.Brit.Psychol.Soc.(1959) pp 38-39.

LAKOWSKI, R. Is the deterioration of colour vision with age due to lens or retinal changes. Tagungs Bericht International Farbtagung Dusseldorf 1961- Teil 2 Die Farbe 11 Nr. 1/6 pp 69-89. (1962).

LAKOWSKI, R. Testing of Colour Vision in Prospective Printers Apprentices and the problems this present in Selection. 6 emes Journees Internationales de la Couleur Evian 1962 (in print).

LE GRAND, Y. Light, Colour and Vision. Chapman & Hall Ltd., 37 Essex Street, London W.C.2. (1957).

LUCKIESH, M & I.W. MOELLER. Naming the common signal colors. J.Opt.Soc. Amer. (1926) 13,pp 465-469.

McDOUGALL, W. An investigation of the color sense of two infants. Brit.J. Psychol.(1906-8) 2, pp 338-352.

MONROE, W.S. Color sense of young children. Paidologist, (1907) 9,pp 7-10.

MORLEY, T. Colour vision. Picture Post. May 12th. (1951).

MYERS, C.S. Some observations on the development of color sense. Brit. J.Psychol. (1908) 2, pp 353-362.

NEWHALL, S.M., D. NICKERSON & D.B. JUDD. Final Report of the O.S.A. Subcommittee on the Spacing of the Munsell Colours. J.Opt.Soc.Amer.33 (1943) p. 385-418.

NICKERSON, D. Tables for use in Computing Small Colour Differences. American Dyestuff Reporter, August (1950).

NICKERSON, D. & K.F. STULTZ. Color Tolerance Specification. J.Optical Soc. Amer. (1944) 34, pp 550-570.

OSTWALD W. Farbkunde, S. Hirsel, Leipzig (1923).

OSTWALD W. Color Science. Part II (Translated by J. Scott Taylor), Windsor & Newton, London, (1933) pp 35-46.

PETERS, G.A. The New Dvorine Color Perception Test. The Optometric Weekly (1954).

PHILIP, B.R. The discrimination of color mass. Amer.J.Psychol.(1938), 51, pp 127-135.

PICKFORD, R.W.. Individual Differences in Colour Vision. Routledge and Kegan Paul Ltd., London, 1949.

PICKFORD, R.W. An Item-Analysis of the Ishihara Test. Brit.J.Psychol.(Gen. Sec.) (1950) Vol.XLI.

PIERCE, W.O'D. Individual Differences in Normal Colour Vision. (A Survey of Recent Experimental Work (1910-31) Med.Res.Council, Rep Committee. Physiol Vision, London. (1933).

PITT, F.H.G. Characteristics of dichromatic vision. Med.Res.Council Rep. Committee Physiol. Vision XIV; Spec.Red.Ser.No.200, London (1935).

PRATT, K.C., A.K. NELSON & K.H. SUN. The Behavior of the Newborn Infant. Columbus; Ohio State Univ.Press. (1930).

PURDY, D.McL. On the Saturation and Chromatic Thresholds of the Spectral Colours. Brit.J.Psycho.21 (1931) p 283.

QUELLETTE, L.A. Age Differences in Colour Discrimination, Unpublished Masters Thesis, Fordham University. (1955).

ROBERTSON, G.B. Colorimetric Analysis and standardisation of Three Tests of Colour vision. Ungrad.Thesis, Univ.Edinburgh, 1964.

ROHNER, E. Die Anwendung des Spectromat in der Textilveredlungsindustrie. Textilindustrie Heft 4. 1963. pp 269-276.

ROSENBERG. Zs.F.Sinnesp. 59,103.(1928).

RYDIN, H. De la capacite de distinguer, aux differents ages de la vie, les radiations vertes ayant une longeur d'onde de 539uu et celles vert-bleuatre de 500uu. Comptes rendus Soc.Biol. (1927) 96,pp 818-820.

SAID, F.S. & R.A. WEALE. The Variation with Age of the Spectral Transmissivity of the Living Human Crystalline Lens. Gerontologia 3 (4) : 213-231 (1959).

SCHOFIELD, R.K. The Lovibond Tintometer adapted by means of the Rothamsted device to measure colours on the CIE system. J.Sci.Instr.16.p74 (1939).

SCHMIDT & HAENSCH. Manual on the Nagel Anomaloscope Mod II. Schmidt & Haensch, Berlin.

SHERIDAN, M & A.D. SORSBY. Refraction and Its Components during the Growth of the Eye from the Age of Three. Med. Res.Council Special Report No.301. 1961.

SIEGEL. Nonparametric Statistics : For the Behavioral Sciences. McGraw-Hill Book Co.,Inc. (1956).

SLOAN, L.L. The Use of Pseudo-Isochromatic Charts in Detecting Central Scotomas due to lesions in the Conducting Pathways. Amer.J.Opth. 25,(11) (1942).

SLOAN, L.L. An Improved Screening Test for Red-Green Color Deficiency Composed Available Pseudo-Isochromatic Plates. J.Optical Soc.of Amer.(1945) Vol35. No.12. pp 761-766.

SLOAN, L.L. Comparison of the Nagel Anomaloscope and a Dichroic Filter Anomaloscope. J.Opt.Soc.Amer. 40 (1950) pp 41-47.

SLOAN, L.L.& A. HABEL. Tests for Color Deficiency Based on the Pseudoisochromatic Principle. A.M.A. Archives of Opthal. (1956) Vol.55. pp 229-239.

SPECIALE P. & F. BOZZONI. Senso Cromatic & Senilita. Boll.Ocul. 40,(1961). 785.

STAPLES, R. Color Vision and Color Preferences in infancy and childhood. Psychol.Bull. (1931) 28, pp 297-308.

THOMSON, L.C. & W.D. WRIGHT. The Convergence of the Tritanopic Confusion loci and the Derivation of the Fundamental Response Functions. J.Opt.Soc.Amer. 43, (10) 1953.

TIFFIN, J. Industrial Psychology (2nd Ed) New York, Prentice-Hall,Inc.(1947).

TIFFIN, J & N.S. KUHN. Colour discrimination in industry. Arch.Opthalm.28. (1942) pp 851-859.

TINTOMETER. Directions for the Assembly and use of the Lovibond-Schofield Tintometer (Tyne 1A). Tintometer, Salisbury, England.

TOKYO MEDICAL COLLEGE. Color Vision Test. Murakami Color Research Laboratory, 10-7 Nishmaka-Dori, Tsukizima, Chuo-ku, Tokyo, Japan.

TRENDELBURG. Zur Diagnostik des abnormen Farbensinnes. Klin.Mbl.Augenhk., 83, (1929) pp 721-732.

TUCKER, A.W. Observations on the color vision of school children. Brit.J. Psychol.(1911) 4, pp 33-43.

VALENTINE, C.W. The colour perception and colour preference of an infant. Brit.J. Psychol. (1914) 6, pp 363-383.

VERNON, P.E. & A. STRAKER. Distribution of Colour-Blind Men in Great Britain. Nature (1943) Vol.152.p 690.

VERRIEST, G. Further studies on Acquired Deficiency of Colour Discrimination. J. Opt.Soc.Amer. 53 (1) (1963).

VERRIEST, G., R. VANDEVYVERE et R. VANDERDONCK. Nouvelles Recherches se Rapportant a l'Influence du Sexe et de L'Age sur la Discrimination Chromatique Ainsi Qu'a la Signification Pratique des Resultants du Test 100-Hue de Farnsworth-Munsell. Revue d'Optique. (1962) pp 499-509.

WALRAVEN, P.L., H.J. LEEBEEK & M.A. BOUMAN. The I.S.C.C. Color Aptitude Test : The Interpretation of some testing results. Instituut Voor Zintuigfysiologie R. V.O. - T.V.O. 1956-

WARBURTON, F.L. Variations in normal colour vision in Relation to Practical Colour Matching. Proceedings of the Physical Soc. LXV11,6-B(1953) pp 477-484.

WEALE, R.A. The Aging Eye. H.K. Lewis & Co.,Ltd. London 1963.

WEAVER, K.S. A Provisional Standard Observer for Low level Photometry. J.Opt. Soc. Amer. 39, (4) 1949.

WILLIS M.P., & D. FARNSWORTH. Comparative Evaluation of Anomaloscopes. Med. Res. Lab.Rep No.190 Bur. Med. and Surgery., Navy Dept. (1952).

WINCH, W.H. Colour names of English school children. Amer.J.Psychol.(1910),21 pp 453-482.

WRIGHT, W.D. The Characteristics of Tritanopia. J.Optical Soc.Amer.(1952) Vol.42. pp 509-521.

WRIGHT,W.D. Researches in Normal and Defective Colour Vision (London : Henry Kimpton). (1946).

WRIGHT, W.D. The measurement of colour. London,Hilger, 1944.

ZANEN, J. Introduction a l'Etude des Dyschromatopsies Retiniennes Centrales Acquises. Soc.Belge d'ophtal. (1953).